Musical reuse in the Notre Dame repertory: historiography and new computational directions

Joshua J. Stutter

Submitted in fulfilment of the requirements for the
Degree of Doctor of Philosophy

Supervisors: Prof. David McGuinness, Dr Tim Duguid, and Dr Joanna Tucker

University of Glasgow

May 2025

Abstract

Of the many questions that we may wish to ask of the repertory of twelfth- and thirteenth-century music commonly known as Notre Dame polyphony, issues surrounding musical reuse and borrowing suffuse nearly every aspect. Questions of chronologies, locations, composers, purposes, processes, orality, performance, transmission, as well as all manner of other unknowns in the repertory, can be well argued and evidenced by the presence (or not) of musical reuse and borrowing. The freedom with which fragments, and sometimes even entire settings of music, circulated throughout thirteenth-century musical culture appears to defy our typical notions of chronology, composition, and even deterministic process.

Musicological scholarship from the 19th century to the present day places the clausula at the centre of this whirlwind of borrowings and musical–textual allusions, the enigmatic form by which the contemporary music theorist Anonymous IV contended that the modernisation of the repertory had taken place. Modern study has identified many so-called substitute clausulae in the fragments of polyphonic settings transmitted in certain fascicles of the central manuscript sources, and it is largely by these substitute clausulae that the later genre of the motet was created. However, relating substitute clausulae back to their source organa and forward to motet has proven difficult, due to the music of the Notre Dame repertory refusing to yield in all cases to the simple and causal chronology that this implies.

More recent work in particular is considering how musical reuse in the Notre Dame repertory may be conceived beyond the static and one-to-one substitutions of clausulae. These are often thought of as different phenomena: processes such as indirect concordances and melodic formulae are considered entirely separate from clausula substitution. At the same time, musical reuse has been treated with a particularly restrictive historical focus, and the clausula has been given centre stage as the main vehicle for musical reuse, catalogued and analysed as tightly interwoven with motet, and minimising other forms of musical reuse in the process. However, when we try to uncover what authors actually mean when they are discussing aspects of musical reuse such as the clausula, we find their definitions inexact and conflicting.

This dissertation therefore tackles the issue of musical reuse in the Notre Dame repertory afresh, placing the notation of the manuscript sources at the centre of the study and considering how the broader issue of reuse can be studied holistically in the twenty-first century, beyond limiting categorical descriptors and arbitrary criteria. By creating new methodological tools around the complex and ambiguous notation of thirteenth century polyphony, I investigate how new digital methods can be harnessed to

extract patterns of musical reuse further than via simple verbatim substitution. The operation and interpretation of thirteenth century notation poses new and unique challenges for representation in the computer which have not yet been satisfactorily solved. I therefore develop novel encoding, optical music recognition, and analysis methodologies to approach these problems from new angles, centred around a specially created online database for the browsing, editing, and analysis of thirteenth-century polyphony.

Using the novel tools that I present as part of this study, I extract new and previously undetected patterns of musical reuse within the repertory that fall outside of our typical perceptions of independent reuse phenomena, pointing the way towards a wider and more nuanced concept of musical reuse as a dispersed network of common musical culture in the twelfth and thirteenth centuries. This is a concept that does not necessarily rely on causality or linear process, and hence may resolve the underlying issues (chronology, composition, and deterministic process) that musical reuse represents.

Contents

List of Figures

List of Tables

Acknowledgements

My introduction to the topic of Notre Dame polyphony was through the final-year undergraduate elective “Parisian Polyphony” led by Susan Rankin and Sean Curran. This is now the third PhD presumably sparked from that one course in 2014–2015. In that same year, much of my current thinking also began in the long breakfasts shared with Nick Doig: we would frequently bemoan the dry and precious manner in which much early music is often performed, and effuse over the few recordings we could find where early music was tackled with new and vivid approaches. A hypothetical vocal ensemble was imagined as antithesis to all we saw as sterile and lackluster in performance practice, not afraid to perform medieval music “low, fast, and loud”.

In 2017, I founded *iuchair*, partly with these goals in mind, and the music of the thirteenth century only dimly in view. It is almost incredible to me now that Edward Marshall, Alasdair Robertson, and Harold Thalange responded positively to an extremely strange, out-of-the-blue e-mail from a perfect stranger with nothing more than a vague plan to get together and perform some thirteenth-century organum. However, respond they did, and I can distinctly recall in those first rehearsals the editions that I had naïvely created as “ideal” for performance were instantly set upon, scribbled on, tore up, and pasted over as we struggled to come to terms with the immense difficulties that arose as we attempted to invent a new performance practice. The first performances of *iuchair* — the name first a nonsensical joke stemming from a *Limmy’s show* sketch, then serious Gaelic loanword, then a joke again — then became a question of what repertoire to tackle next. Six years later, the continued input of the members of the ensemble – since then also most notably Edward Horrocks and Joshua McCullough — have time and time again challenged my opinions of how early music “ought to sound” and the relationship of academia to performance as more than just a one-way street.

It was also in 2017 that Katy L. Cooper wisely informed me that the work I was doing would be better framed in the context of academic study, and I first met with Andrew Bull who had recently completed a compelling Master’s thesis at the University of Glasgow on the *conductus* of W_1. It was through this circuitous route that I decided to return to academic study, first through a Master’s on the impact of oral culture on the transmission of W_1, and now this doctoral study. This study, which commenced in October 2019, has of course been significantly affected by the COVID pandemic, not necessarily in terms of substance, but in the opportunities that have been afforded for discussion with other scholars. I was grateful to have an early discussion with Greta Mary Hair, but since then my thinking has been shaped by conversations I have had online and in person with: Warwick Edwards, Stefan Morent (and the Tübingen *eChant* working group to whom I presented my work), Michael Friebel, as well as Giuliano di Bacco & Anna Plaksin (and the MEI-Mensural interest group). I have also benefited from e-mail correspondences with (but not limited to): Catherine A. Bradley (on defining clausulae), Wulf Arlt (on his extensive research in the area), and Henry Parkes (some sleuthing related to previous digital projects in this area at the University of Nottingham).

Earlier stages of this work have also been presented at various conferences, where anonymous reviewers as well as conference attendees have provided invaluable feedback: the Music Encoding Conference 2021 (online), Digital Libraries for Musicology 2021 (online), MedRen 2023 (Munich), the joint MEC/TEI conference 2023 (Paderborn), *The Future of Digital Editions of Music* (David Smith and Frauke Jürgensen, Northumbria University), and *Benedicamus Domino* (Catherine A. Bradley and team, University of Oslo).

Others have also read and provided feedback on this work at various stages of its creation: Eva Moreda Rodriguez, Yunhyong Kim, and Alice Masterson. Above all these however is my father, Dan Stutter, who has proofread every word in this dissertation countless times from an outsider perspective (but any further typos are my own!) Many ideas were also written as part of my blog, and I am grateful to the many e-mails I have received from around the world from people interested in my work or pointing to further avenues of research. This dissertation was typeset with the help of Peter Schaffter's excellent "Mom" macros for GNU Troff, and graphs using Ted Faber's port of "grap".

On a more personal level, my thanks also go to my parents and Sam for their continued support; Edward Marshall for his linguistic perspective and often acting as sounding board for some of my stranger ideas; Clara Miller for being endlessly supportive (especially during the soul-destroying "clicking" portion of my research); and finally to Eilidh who not only helped me clean up some blurry-looking figures, but also provided love and support, quick to offer congratulations and consolations during the repeated genius–idiot mood cycle.

Author's declaration

"I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution."

Joshua Stutter

Definitions and Abbreviations

ACID	Atomicity, Consistency, Isolation, Durability
AJAX	Asynchronous JavaScript and XML
API	Application programming interface
AUROC	Area under the receiver operating characteristic curve
CANDR	The Clausula Archive of the Notre Dame Repertory
CBOW	Continuous bag-of-words
CNN	Convolutional neural network
CRUD	Create, Read, Update, Delete
CSS	Cascading Style Sheets
CSV	Comma Separated Values
CUDA	Compute unified device architecture
CWMN	Common Western music notation
DIAMM	Digital Image Archive of Medieval Music
DTW	Dynamic time warping
EMD	Earth movers' distance
ERD	Entity relationship diagram
F	Biblioteca Medicea-Laurenziana, Florence. MS Pluteus 29.1
FAIR	Findability, Accessibility, Interoperability, and Reusability
GCN	Graph convolutional network
GNN	Graph neural network
GPU	Graphics processing unit
GUI	Graphical user interface
HAB	Herzog August Bibliothek
HMM	Hidden Markov model
HTTP	Hypertext Transfer Protocol
IIIF	International Image Interoperability Framework
JSON	JavaScript Object Notation
KUR	Key user requirements
LSTM	Long short-term memory
MEI	Music Encoding Initiative
MGG	Musik in Geschichte und Gegenwart
MoSCoW	Must, should, could, wish

MüA	Bayerische Staatsbibliothek, Munich. Mus. MS 4775
MVC	Model–View–Controller
ND	Notre Dame
NLP	Natural language processing
OMR	Optical music recognition
ORM	Object relational mapper
OT	Optimal transport
PCA	Principal component analysis
RDBMS	Relational Database Management System
RDF	Resource Description Framework
RDFa	Resource Description Framework in Attributes
RNN	Recurrent neural network
ROC	Receiver operating characteristic
RoPE	Rotary Positional Encoding
SAX	Simple API for XML
SQL	Structured Query Language
URI	Uniform Resource Identifier
VGG	Visual Geometry Group
W_1	Herzog August Bibliothek, Wolfenüttel. Cod. Guelf. 628 Helmst.
W_2	Herzog August Bibliothek, Wolfenüttel. Cod. Guelf. 1099 Helmst.
XML	Extensible Markup Language

- **Pitches** are indicated using Scientific pitch notation
- **Manuscripts** are cited by their sigla (given above) in bold type
- Where appropriate, settings of music are also cited by their **Ludwig number**
- The citations of particular **systems** of music are indicated by upper case roman numerals following the folio number
- **Colours** used in plots follow two colour schemes. Qualitative categories (such as types) where the number of categories is 20 or fewer follow the **tab20** colour map.[1] Sequential data, or where the number of categories is greater than 20, follow the **viridis** colour map.[2]

[1] See <https://matplotlib.org/stable/users/explain/colors/colormaps.html#qualitative>.
[2] See <https://sjmgarnier.github.io/viridis/>.

Chapter 1
Introduction

The collection of twelfth- and thirteenth-century music associated with the cathedral of Notre Dame of Paris is commonly termed "Notre Dame" (ND) polyphony.[1] It is a repertory and associated academic discourse that appears to revolve around its central contradictions, being on the one hand perhaps the most well known medieval repertory due to its historiographic impact on what is now considered Western music, and yet there being so little understood about it. As a result, it is one of the first styles of polyphonic music to be mentioned in histories of Western music,[2] but it is also one of the most difficult of which to gain a complete understanding.[3] It is generally marked out as a leap from monophony to polyphony and from oral to literate musical culture in the Western tradition,[4] but it was not the first polyphonic practice in Western Europe nor is it known to what extent literacy was part of its creation (Fuller 2011). Its title is an eponym for Notre Dame of Paris in the thirteenth century, but the term ND repertory (or school) is also used as a catch-all term for a vast swath of thirteenth-century music, especially given the repertory's links to genres such as motet through clausulae.[5] Not only is the repertory frequently held up as 'a paradigm shift in [Western] music history' (Bradley 2018b, p.1), even with basic questions concerning its creation, preservation and dissemination as yet unanswered, but it also manages to occupy a pluralistic space in which multiple theories surrounding these questions co-exist without necessarily being in conflict with each other. Some rudiments have been accepted for a hundred years or more but other key issues are far from decided, with no apparent

[1] The shortening of the term "Notre Dame" to ND is continued here after the arguments that I made in Stutter (2020, p.3), and this acronym is continued throughout this dissertation. Although this had previously been done by Handschin (1924), it is clear that in his case shortening "Notre Dame" to ND was done simply as a shorthand (there are fifteen other acronyms in the cited article alone). Rather than simple abbreviation, the acronym "ND" as used here aims instead to dissociate "Notre Dame polyphony" from the geographically limiting "Notre Dame", after the arguments made by Losseff (1994, pp.3–24).

[2] 'It would be difficult to overstate the historical significance of the polyphonic tradition cultivated in Paris during the latter part of the twelfth century and for much of the thirteenth' (Roesner 2018, p.835).

[3] For example, Treitler (2003c, p.48), reflecting on the issue of orality in ND polyphony, wondered how it was possible that 'knowledgeable scholars' can have 'so sharp and fundamental a disagreement' on such a key aspect of the repertory. Similarly, Yudkin (2005, p.709) remarked, in his review of one of the more recent attempts to edit the repertory in full, that even a collection of some of the best scholars in the field could not agree on a single editorial practice.

[4] 'It is generally accepted that the Parisian repertory created by Leoninus and Perotinus represents the beginning of polyphonic composition in the modern sense, and indeed of the European polyphonic language itself […] it is a repertory generally assumed to have been conceived and disseminated in writing, not ad hoc in performance, and to have had a stable transmission' (Roesner 2001, p.232).

[5] A note on italicisation: in this dissertation, I have elected not to italicise genre indications such as clausula/clausulae, conductus, copula, discantus, motet, organum/organa, or purum because they occur so often as to become fully adopted words. Italicising each one would cause many parts of the prose to rapidly shift back and forth between roman and italic type, and therefore become distracting and difficult to read.

incompatibility with what are considered the already-established fundamentals.

When confronted with core questions surrounding the repertory, often the best answer is one of pure equivocation. The answers to even basic questions about the repertory require a balanced, nuanced, and occasionally prevaricating response instead of the clear and concise replies that one would think would be requisite to a full understanding of the repertory. Unanswered questions are as basic as:

Who was responsible for the creation of the ND repertory? For every claim towards the identities of those purportedly responsible (e.g. Wright 1986), there is an opposing argument that Leoninus and Perotinus may simply be “projection figures” and, although they may in fact relate to real historical people, they should be evoked merely as representative of changes within the repertory rather than the persons directly responsible.[6]

Why was it written down? Against evidence that the sources of ND polyphony were the direct literate and living results of an already rich musical heritage — conceived either in writing (Roesner 2001, p.232) or orally (Busse Berger 1996) — there is other evidence to show that some sources were copied at arm’s length from the musicians in specialised manuscript workshops (Bevilacqua et al 2018).

What was the intended use of the sources? Arguments that rely on the ‘museum-like’ preservation of some organa (Bradley 2018b, p.249) must contend with some of the possible practical and pedagogical uses that these sources may have had (Treitler 2003a, pp.68–83; .

How was the music disseminated across medieval Europe? Competing theories for the Scottish source $\mathbf{W_1}$, for example, attribute the transmission of its contents to, alternately, a gradual dissemination of the music across Britain (Roesner 1976, pp.378–380), a close working relationship between Paris and St Andrews in the early thirteenth century (Everist 1990), or simply the work of one particularly diligent scribe who collected as much material as was available to him (Steiner 2021, p.397).

After over one hundred years of research and considering the lack of any hard and fast answers, in many fields of study this would be a shameful admission: i.e. that we are arguably no closer to concluding the discussion of these questions, or establishing solid foundations from which to continue in our enquiry.

Perhaps we have been asking the wrong questions. Indeed, in many ways some of the most successful work done recently has been in fact actively seeking to undo what is

[6] ‘There has been ongoing discussion not only about the chronology but increasingly also about their conceivable relationships with each other and about what is attributed to them as their musical legacy. The conviction is beginning to prevail that they are, to a certain extent, projection figures (in the common sense: merely representatives)’ (*Nicht nur über die Chronologie war immer wieder diskutiert worden, sondern zunehmend auch über ihre denkbaren Beziehungen zu einander sowie über das, was man ihnen als musikalische Hinterlassenschaft zuschrieb. Durchzusetzen beginnt sich die Überzeugung, daß es sich zu einem gewissen Grad um Projektionsfiguren (d.h. im landläufigen Sinn: bloß um Repräsentanten) handelt*) (Flotzinger 2007, p.13). See also Haas (1997, pp.871–872).

now perceived as the somewhat rash conclusions of earlier scholars, who unfairly compared the work of Perotinus to Bach and Beethoven (Busse Berger 2002b, p.94), and aimed to assemble various disparate styles from over hundreds of years into grand narratives of national music (Leech-Wilkinson 2007, pp.168–169). This more recent academic culture, of critiquing conclusions that were previously held as gospel, has fostered an emerging concept of medieval repertories such as ND polyphony that considers at its forefront the historiographic and modern pressures on medieval music as key to dispelling past and present tendencies to mythologise the music of this period.[7] However, there is a certain comfort and welcome in this inconsistency; a knowledge that as long as an argument remains consistent with the few basic facts that are settled and can reason well against the considerations of the field, it can be ushered in to join the many other collected theories of the ND repertory, each brandishing its own set of persuasive — but not yet completely convincing — evidence.

This dissertation does not aim to answer any of the above questions directly, but seeks instead to inform these unanswered questions at a more basic level by providing groundwork for how such questions may in future be answered, through the development and demonstration of new methodologies. These will be focused on two angles that I believe underlie these broader questions. First and foremost is the problem of data. In many areas of musicological inquiry, particularly those of current and/or underexplored repertories, musicologists are able to freely define the parameters of their data collection methodologies: music and sources deserving of full-length studies that are yet to be catalogued and studied, practices that are still ongoing which continuously create new evidence to be observed and analysed. What luxury! Although scholars of non-historical musics inevitably have their own issues within their respective fields, one aspect that they are much more in control of is the method and rigour by which they collect their primary data. If a researcher perceives that the current data available to them is inadequate, it is often possible and indeed expected that they should design experiments to collect suitable data directly from the musical objects of study, e.g. composers and performances. In this way, exact hypotheses about the music can be tested, and theories directly proven or disproven by a suitable application of the scientific method.

Study of the ND repertory on the other hand, like many other historical music studies, remains unfortunately unencumbered by new primary data and, rather than being able to collect new data, scholars are forced to parsimoniously review mainly the same

[7] As an example of this subtle mythologising tendency, Treitler (2003a, p.vii) confessed that his early efforts to naïvely discover 'the origin of the conductus' were 'hardly justified by medieval usage or practice, and that this can give a measure of how much invention has been entailed in the construction of that panorama and altogether of the panorama of "Middle Ages", is something of which my awareness was still well in the future at that time'.

music and sources that those first scholars catalogued over a century ago. This is not to say that there is no hope for any progress within the field: progress is consistently made through further analysis of the extant sources. Moreover, since the "cultural turn" of the late 20th century there has been a distinct move in historical musicology away from the dominance of primary source study, what Kerman (1985, p.12) characterised as 'the verifiable, the analysable, the positivistic', and a move towards the synthesis of a broader scope of sources in order to understand better 'music as aesthetic experience'.

A pioneering example of this shift towards culture in the study of the ND repertory is Wright's (1989) study of the musical life of the cathedral of Notre Dame de Paris in the medieval period. Despite not directly introducing any new sources of musical notation to his study, Wright managed to make large advances in the understanding of music at Notre Dame by considering sources other than musical notation and the writings of music theorists. This valuable cultural work is continued in studies that consider the people and cultures that created these musical artefacts through other traces than just musical notation and music theory (e.g. Page 1989, pp.134–154). Most recently, an article by Nemarich (2024) has furthered understanding of the makers and making of organum, not by scrutinising the minute detail of **F**, $\mathbf{W_1}$ or $\mathbf{W_2}$, but by considering how these musicians may have interacted with merchants, secular musicians, and the clergy within a wider cultural sphere.

Despite these advances, what cannot be done with either primary source study or cultural study is to collect any new data from the repertory that can be controlled directly. The sources that we have access to do contain data about the musical objects of study, but this musical data has been filtered through the intentions and understandings of the scribes and cultures that interacted with them, and it is thus uncommon for the sources to provide direct evidence for the key questions that we have. As Page (1993, p.11) has observed, medieval music theorists' concerns lie almost exclusively with theory: they seldom betray much information on how that music operates in practice.[8] Within the notational sources, confounding factors such as scribal practice or the intended use of manuscripts can rarely be completely controlled for in any rigorous manner. What is presented in the sources available to us is largely all that we can ever hope to work with in terms of notation, but at the same time can only represent a small slice of the full historical picture (Tosh 1999, pp.168–169).

What we can control for, however, is the way in which we interpret that data. Again, in this respect the ND repertory has been subject to an unfortunate treatment. The dominant role that ND polyphony has typically played in historical narratives of thirteenth-century music, as the herald of new compositional styles and the birthplace of

[8] Although they do admit that the performance of organum was likely much more irregular than the rhythmic modes imply. See the discussion in Roesner (1979, pp. 185–189).

the motet (see, for example, Hoppin 1978, p.215, 252), has in the past unfairly elevated the repertory to a difficult historiographical position where it has been expected to comply with narratives centred on constructing a cohesive sense of "Western art music". Losseff (1994, p.xviii) has termed this bias towards the ND repertory as progenitor of Western art music the 'hegemony of Notre Dame', and Page (1993, pp.1–42) has also critiqued the tendency of historians to monumentalise medieval polyphony by exaggerating its "architectonic" characteristics.

Whereas Page (1993, p.5) attributed this inclination to a broad sense of 'idealism' endemic to medieval historians, Losseff (1994, p.xviii) and Busse Berger (2005, pp.9–44) have laid the blame for this difficult historiographical status squarely at the feet of the work of the 'founding father of medieval music' (Busse Berger 2005, p.9), Friedrich Ludwig, and his enduring legacy. This is especially the case for Ludwig's *Repertorium Organum Recentioris et Motetorum Vetustissimi Stili* (1910), over a century later still a core text in ND studies. Although Ludwig's conclusions of musical "progress" from organa to motet via the substitute clausula are now generally viewed with a healthy dose of scepticism, the true picture is often obscured by his simplistic linear narratives that have since percolated through academic literature largely unchallenged, especially in the case of motet (Bradley 2013, pp.2–3), the focus of Ludwig's study. This emphasis on a narrative that culminates in Latin motet permeates Ludwig's work, where even his numbering systems can be seen to have 'enshrined' certain biases (Bradley 2018b, p.5).[9]

However, if we focus solely on his materials it is difficult to argue that Ludwig's conclusions in *Repertorium* would have been any different if he had had access to the somewhat wider range of sources we have at our disposal today. The shortcomings of Ludwig's work, as well as other positivistic scholars working in the late nineteenth and early twentieth century, are not in the data that was available to them, but in their interpretative perspectives and categorical methodological models. Modern perspectives on the ND repertory are doubtless an improvement, in large part due to the cultural turn which has granted a wider cultural context in the treatment of medieval polyphony. Nonetheless, this change in perspective has not been accompanied by a larger change in methodology to reflect this broadened scope. Indeed, modern scholarship uses essentially the same methodological tools that scholars in the early twentieth century had at their disposal for the analysis of ND polyphony, focusing on 'close and critical reading' (Bradley 2018b, p.5), albeit geared more towards a 'case-by-case approach' and placed within a suitable cultural context rather than constructing 'developmental' narratives

[9] See also Tischler (1984, p.166): 'we cannot speak of organa as wholes and even less as single items related to particular chants, as the order numbers introduced by Friedrich Ludwig would suggest and as Husmann treats them. Thus Ludwig's O1 or M1, for example, not only represent three different settings each, but each is apparently the result of several hands or acts of composition'.

(Bradley 2018b, p.9).

Despite this paucity of new material to study, analysis of the sources available to us is by no means complete. Far from there being no more data to extract from the sources for study, very little of the notation within the manuscript sources has been analysed in detail. The individual movement of notes within voices, ligature patterns, and even the layout of the folios are veritable treasure troves of data for the study of ND polyphony of which studies such as Ludwig's only scratched the surface. It could very well be possible that answers to some or all of the open questions in the field lie not in the simple cataloguing and concordance of polyphony, such as was first attempted in Ludwig's *Repertorium* (1910), but in a deeper and more fine-grained look at the minutiae of the repertory's form, style, and notation. This study therefore emphasises a return to the central and primary sources of the repertory: **F**, $\mathbf{W_1}$, and $\mathbf{W_2}$, with a view to investigating how much more data can be extracted directly from these incredibly rich manuscripts.

The second angle that will be considered in this study is musical reuse, a facet of the repertory that lies at the heart of many of the open questions listed above. ND polyphony has attracted so much study not only because of its supposed 'epoch-making' developments (Roesner 2001, p.232), but also due to the liberty with which it borrows and circulates its musical material. It is not uncommon to see transmissions of organa substitute passages of polyphony between each other, especially discant sections, as "substitute clausulae" (Flotzinger 1980). This is a key part of musical reuse in ND polyphony, not least due to the fascicles of $\mathbf{W_1}$ and **F** that are dedicated to transmitting substitute clausulae, but also the fact that such clausulae were often one of the main sources for early Latin motet. This process has been known for over a century (Meyer 1898), and there are numerous studies that have dealt with the substitute clausula directly.[10]

Beyond this well known aspect of musical reuse however, there is a second and much less frequently studied level of musical reuse, particularly within the cycles of two-part organa, of concordances between settings of polyphony that do not appear to be so purposefully linked. The easiest of these to detect are what Payne (1996, p.338) has termed "indirect" concordances: passages where part of a duplum bears almost exact similarity to what at first may seem an unrelated passage, but the settings are not otherwise related by tenor melody. Whereas clausula concordances within the ND repertory have been catalogued in full and in detail — beginning with Ludwig (1910) and updated most recently by van der Werf (1989) — indirect concordances within the

[10]For an up-to-date account on the history of the study of substitute clausulae, see Mathias (2020), especially pp.12–20. For the clausula's impact on early motet, see Bradley et al (2019, §I. Middle Ages, 1. Thirteenth-century origins). The particular historiographic issues surrounding the use of the term "clausula" will be addressed more fully in chapter 2.

repertory have not yet been treated to the same level of scrutiny, and this is perhaps due to the difficulty of detecting them comprehensively.

Although Payne (1996, pp.338–402) has begun this painstaking process for the indirect concordances within the organa of $\mathbf{W_2}$ in his edition of the manuscript, detection of indirect concordances has been done most fully by Tischler (1988, pp.39–77) in the two-part organa, where he catalogued his interpretation of each setting's "interrelationships", as he saw them. However, by his own admission, these analyses are 'somewhat arbitrary' (Tischler 1988, p.39), and his observations rarely amount to much more than remarking that certain settings 'may lean on' others, or were 'mined' for material (Tischler 1988, p.41). Neither Tischler nor Payne considered these indirect concordances for collation in the edition. There may very well be other indirect concordances that have escaped attention, and we have no way of verifying these concordances without checking the sources ourselves to see if they match our own arbitrary criteria.

As an example of this work that is yet to be done, we can briefly consider how these indirect concordances may transform our understanding of a *unica* within the repertory. *Alleluya. Benedictus es* (M57) is transmitted only once within the entire repertory (**F**, f.143v–144r) and it has been catalogued as containing no clausula concordances (van der Werf 1989, p.88). It therefore must be considered independently of other settings, and very little can be inferred about its transmission. However, an indirect concordance not catalogued in any existing study tells us something about its relationship to the grand, four-part settings of the repertory. The passage on "patrum" (f.144r, I) comprises two small snatches of polyphony from both the four-part *Uiderunt omnes* (M1, transmitted in **F**, f.1r–4r) and the four-part *Sederunt principes* (M3, transmitted in **F**, f.4r–7v). The first two *ordines* of "patrum" are identical to the first voice of f.6v, III ("domine"), and the third and fourth *ordines* are very similar to the third voice of f.1v, I ("uiderunt") (Fig. 1.1).

Far from *Alleluya. Benedictus es* not having any concordances, it has links to two of the largest and most developed pieces within the repertory, and it would be remiss for discussion of *Alleluya. Benedictus es* not to consider its possible links to the *quadrupla*. Does M57 represent a later setting that took its cue from four-part discant, or does it simply form part of the same formulaic practice that birthed both it and the two *quadrupla* in question? Further discussion of this indirect concordance is out of scope here, but the general point is that, far from being insignificant details, these subtle links between organa have the potential to provide compelling evidence for musical reuse outside of the typical bounds of the substitute clausula. Crucially however, this indirect concordance is not exact and, although the *Uiderunt* concordance is similar enough to be considered concordant, it ends with a slightly different formula. This brings us to the issue of detecting more subtle concordances within the repertory, and raises the question of

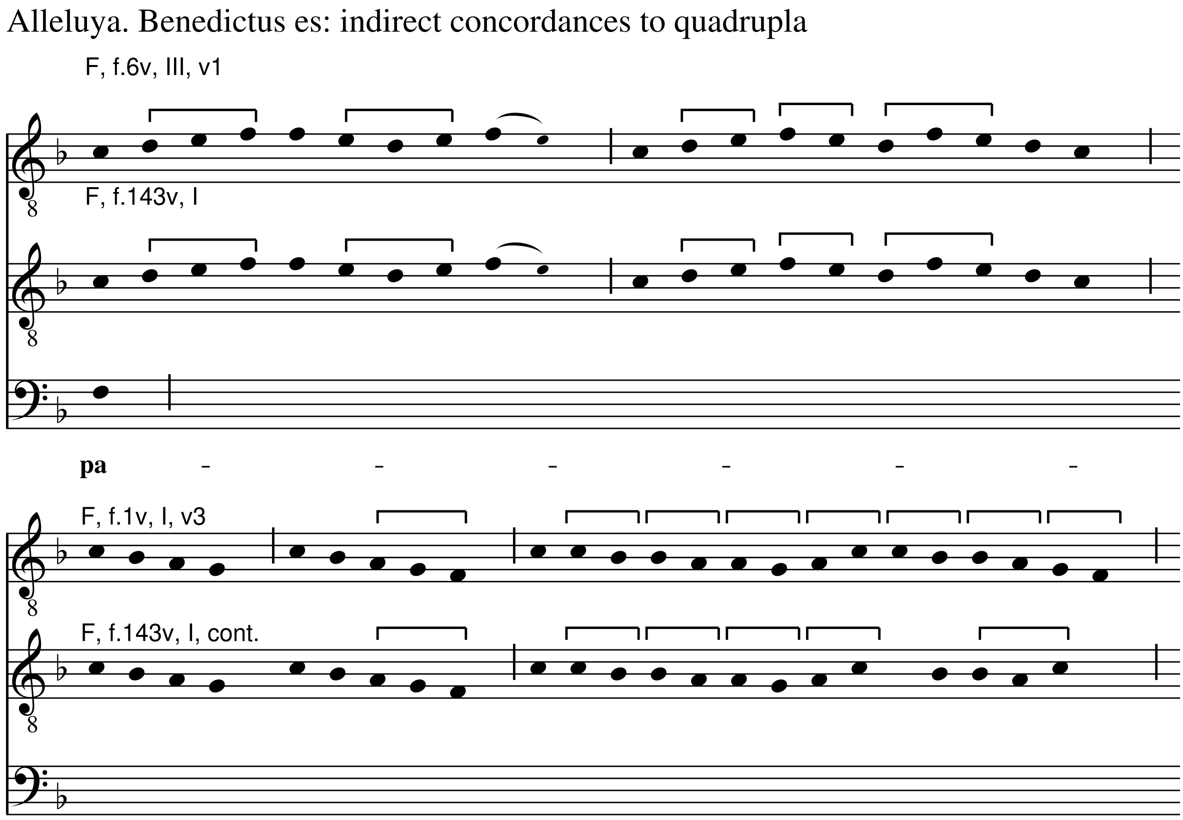

Fig. 1.1

what should be considered a true concordance and what can simply be attributed to coincidence or indeed a common formulaic gesture?[11]

Musical reuse is therefore one of the fundamental issues in the repertory, as it seemingly moves in opposition to the static and compositional directions frequently attributed to ND polyphony. Musical reuse within ND polyphony exists in stark contrast and occasionally in contradiction to these compositional aspects (Roesner 2001, p.234), and so offers a way into some of its more elusive considerations. However, its manifestations beyond the typical concordances of the substitute clausula are understudied, and this is largely due to the mammoth effort that would be required to find and catalogue all indirect concordances and melodic formulae. Indeed the scale of such an apparatus, if incorrectly targeted, could in fact serve to obscure rather than illuminate the processes behind ND polyphony.

Nonetheless, studying musical reuse can allow us to more closely examine the "who", "why", "what", and "how" questions noted above by analysing how the processes behind musical reuse impact or interact with each. Clausulae substitutions largely submit themselves successfully to cataloguing efforts, as the cataloguing model reflects the static, literate, and mostly one-to-one swaps apparent in this form of musical

[11] This issue will be discussed in more detail in "2.10. Musical reuse in the computer".

reuse (Roesner 2001, p.265). However, more subtle musical reuse in the indirect concordances and melodic formulae across the repertory will likely not be able to be catalogued this same way. This is because the model for formulae, perhaps more based on *memoria* than the construction of clausulae (Roth-Burnette 2010, pp.20–29), is inherently more transformative: formulae can be transposed and their configurations altered in any number of ways (Roth-Burnette 2010, p.376). Attempts to catalogue indirect concordances via static critical apparatuses — such as tables, lists, and diagrams of stemmata — are therefore destined to failure. The categorical nature of concordance vs non-concordance assigns this variable into one of two diametrically opposed categories, losing the all-important nuance and shades of grey in between.

This dissertation aims to find new ways into conceiving of musical reuse within the ND repertory. Rather than building upon the existing taxonomy of substitute clausulae, concordances and formulae by aiming to generate comprehensive lists of each, this study conceives of musical reuse as a single continuum where substitute clausulae form only the most obvious and literate layer. This study therefore starts afresh from the sources as they lie, as scholars such as Ludwig did, but aims to use the information contained within the ND sources to generate a breadth of data, as evidence that can then be analysed at once using twenty-first-century technology. Rather than positioning itself within the already-established framing of discourses that attempt to reason with a certain selection of evidence in order to argue for a particular view on a broader question, this dissertation sidesteps those issues to consider primarily the notation of the source manuscripts in the context of the question of musical reuse.

Although close reading continues to provide valuable insights into the analysis of the ND repertory, it becomes difficult to scale that analysis to the consideration of more than a few settings at once. The process of musical reuse is not just a local phenomenon, limited to a select number of settings, but is reflected within the ND repertory at large. Concordances may be found in unexpected places, and the consequences of musical reuse may affect how we view the repertory as a whole. Roesner (1981, p.398) has shown that the philological apparatuses typically used to analyse such large-scale structures, such as Lachmannian stemmatics, are not workable in the context of the ND repertory, and this is exemplified by Tischler's (1988, p.39) own admission that his stemmata for the two-part organa are 'somewhat arbitrary'. At the other end, on the scale of close analysis, Bradley (2018b) has shown that the analysis of thirteenth-century music is indeed possible, but how we might advance this analysis to repertory-scale study is not clear. Furthermore, repertory-scale study that necessarily relies upon what is typical or characteristic within a style lacks transparency in how we define and evidence the 'intuitive statistics' of our own personal conceptions of the repertory (Neuwirth and Rohrmeier 2016, pp.175–176). This research aims to provide solutions to these problems at the larger scale, by investigating new methodological routes for the corpus study of thirteenth-century music. This will be done by looking at musical reuse at the repertory scale rather than close reading, and will harness advances available from more modern and digital methodologies, so-called "distant reading" (Moretti 2013), corpus study, machine learning, and Big Data.

1.1. Research questions

Musical reuse within the ND repertory is most often viewed through the lens of the substitute clausula. For over a century, it has been known that the ND repertory was not a static collection of music but was actually much more fluid, where there are often multiple versions of the same polyphony extant, even within the same manuscript source (Meyer 1898). The so-called substitute clausula is emblematic of this process, where alternate passages of polyphony, most often in discant, are copied in separate sections of the manuscript away from their original context.[12] It is easy to observe in the substitute clausulae the striking way in which these borrowings and substitutions have not been demoted, hidden, or scored out in the manner of precursors or early drafts, but meticulously preserved alongside other versions, presumably for continued use within the repertory. Both contemporary accounts and modern accounts from the nineteenth century to the present day use the term "clausula" to describe some form of musical section, but neither the ancient nor the modern go as far as to draw the parameters of clausulae exactly.

Lists of clausulae have been created before, and the way that sections of polyphony have been substituted and interact between sources through this process has been catalogued. Ludwig (1910) initially identified 57 such clausulae, although he failed to number them consistently (Smith 1964, p.163). Dittmer (1959b, p.126) added 28 more, and Smith (1964, p.13) two of his own, bringing the current total number of clausulae through these studies to 87. What is meant by "clausulae" in these lists however, is not "substitute clausula" but any discrete discant section, whether alternatives exist within the substitute clausulae fascicles or not (Dittmer 1959b, p.102). This definition does not describe musical reuse in all 87 cases, but simply the existence of discant sections within the repertory as opposed to purum, and included in this list are discant sections for which there are no major divergences between manuscripts.

For example, Smith (1964, p.185) includes "sunt" from *Dum complerentur* (O11) in his list of clausulae, despite all three transmissions being essentially identical, and there being no "sunt" substitute clausulae. Conversely, Roesner (2001, p.258) describes

[12] Due to the imprecision in the usage of the unqualified term "clausula" (see "Chapter 2: Historiography of musical reuse: the clausula and its context"), I will use the term "substitute clausula" after Bradley (2015, p.154) to refer specifically to 'passages removed from their organal context and copied alongside other such [substitute] clausulae in a separate section of the manuscript'.

the preceding “omnes” section as a clausula in his discussion despite it being clearly in purum style with a held tenor. “Omnes” is divergent in all three transmissions, but does not appear in Smith’s list as it is not in discant. There is a clear disconnect here between how each study actually defines what constitutes a clausula. Whereas a substitute clausula is clearly defined by its position in the manuscript, the term “clausula” as used in literature is used to denote more than just its substitutes: it is not only transmitted as a substitute but also when it has been substituted.

However, given the clausula’s close proximity to the issues of musical reuse, reuse within the ND repertory has been studied largely through clausula substitution. By comparing the “versions” of polyphony implied by the different substitute clausulae preserved for each setting, scholars have been able to hypothesise the ways in which these substitutions could have operated. Despite musical reuse necessarily forming a core part of thinking related to the clausula, the central studies of clausulae — such as Smith (1964), Flotzinger (1969), Baltzer (1974), and most recently Mathias (2020) — are not studies of musical reuse directly but of the styles of discant composition in the ND repertory. Clausula studies have been generally focused on describing, cataloguing and analysing discant as a new style in ND polyphony rather than investigating clausulae in terms of musical reuse.[13]

Although there have been some writers who believe that the substitute clausulae were not used for substitution at all but were instead independent compositions — most notably Harrison (1958, pp.122–128) — the general consensus is that the key function of these substitute clausulae was to replace corresponding sections in the *Magnus liber* (Baltzer 1995, p.xxxix), and so it remains the case that the main purpose of the substitute clausula was to facilitate musical reuse. This study of musical reuse cannot therefore afford to ignore the substitute clausula entirely as a solved issue in favour of an alternate process such as melodic formulae, but the substitute clausula must fit within the same model as indirect concordances and melodic formulae. Nor do I believe it wise to discuss only one polyphonic style such as purum, or only one slice of the repertory, as styles are so closely pressed together within ND polyphony and organa rarely consist of only one style. There is no reason to believe that the scope of musical reuse within the repertory will be anything but global, and recent research continues to unearth interrelationships between styles that are not so clearly linked, such as the conductus and the *Benedicamus domino* (Everist 2018, pp.199–213). This dissertation therefore aims to discuss clausulae not in their function as discant within and without clausulae fascicles,

[13] Smith’s (1964, p.5) aim was to create a ‘catalogue raisonné of the clausula [discant section] repertory’, Flotzinger’s (1969, p.13) was to examine ‘the expansions and stylistic changes’ (*der Erweiterungen und stilistischen Veränderungen*) of the repertory in discant, Baltzer’s (1974, p.12) ‘the establishment of notational and stylistic norms’, and Mathias’ (2020, p.20) ‘the tools and techniques for making a clausula [discant section]’.

but as one type of vehicle for musical reuse, going beyond musical reuse as manifested in the substitute clausula by delving into deeper and more difficult to detect layers such as indirect concordances and melodic formulae.

1.1.1. Looking beyond the clausula for musical reuse

Beginning with musical reuse as manifested in the clausula then, how do we know that Smith or Ludwig (or indeed any other writer that uses the term) defined clausulae for their use correctly? What would a list of clausulae look like if we took a "musical reuse" rather than "discant" perspective? Considering this question only serves to open up more questions, such as: What other vehicles are there for musical reuse within ND polyphony besides the substitute clausula? To what extent must a passage display similarity to another to qualify as musical reuse? What kind of variances are tolerated between the "same" passages before they become different, unrelated passages? Must these other concordances begin on a section break such as a syllable change or can they fill in melismata? Can musical reuse overlap and how does the process all fit together?

In short, the question that this research asks is: how should musical reuse be defined and how do concepts such as the clausula fit into this broader scope within ND polyphony? The term "clausula" has been well used, but often with a loose definition, and this research aims first to provide clarity on the historiographic usage of the term, as well as to investigate the significance of musical reuse in the thirteenth century beyond like-for-like clausula substitutions. A successful result would be one that would not only allow us to untangle the clausula's complicated history in the context of musical reuse, but also to assess how to approach issues of musical reuse in the future. Examples of this could be: an exhaustive list of clausulae alongside indirect concordances and melodic formulae; limiting and exact criteria for what does and what does not constitute musical reuse within the ND repertory; a framework for reappraising musical reuse; or suitable metrics for measuring musical reuse within ND polyphony.

The hypothesis is that more complex forms of musical reuse than the substitute clausula, such as indirect concordances and melodic formulae, exist within a subtle network of interrelationships and cannot therefore be catalogued and enumerated in the same way that the clausula has been. The substitute clausula represents musical reuse that is manifested in a set number of ways of performing polyphony over a tenor. However, there are instances of more subtle reuse within the sources, and these give us a glimpse into the true infinite possibilities that lie behind. As such, our ways of listing and evaluating them needs to be fundamentally re-assessed.

1.1.2. Use of digital methods

A second dimension of this research is methodological. Given the focus on defining limits and suitable criteria above, this research will be concentrated on tabulation and concordance of evidence, as well as the analysis of these collations. More broadly, this research will investigate novel methodologies for collating and analysing music, an area that has leant heavily upon textual apparatuses to contain its information within the printed page. This has included, but is not limited to: tables of clausulae transformation into motet, tables of tenor melodies and their respective clausulae, lists of sigla that make reference to commonly referred-to manuscripts and pieces of music, indices sorted by various criteria, and internal referential page links.[14] To read about a clausula, for example, most studies in this field necessarily require the reader to jump between volumes, pages, musical notation, facsimile, tables, footnotes, lists and prose using a dizzying array of sigla and abbreviations, making it difficult to extract information from the research output.[15]

This research takes advantage of the transformative nature of the past half-century of technological innovation for methodology and presentation in order to collate and analyse this research, as well as to develop efficient and engaging methods to present its findings as transparently and persuasively as possible. The intention is for the output of this research to be usable not only by academics who have the necessary skills to trawl through volumes of opaque material, but also via the provision of open-access datasets, increased interactivity, and end-user focused tools for undertaking exploratory research. Rather than committing itself almost exclusively to textual apparatus and manual concordance methodologies, presenting findings as tables and two-dimensional diagrams, this research will explore ways in which both new and old processes can be automated, improved, and radically re-thought using twenty-first-century computational techniques in order to undertake fieldwork and analysis that previously would have taken decades. Moreover, this research investigates how its findings can be presented with as much impact as possible using modern digital presentations of research, such as a database or website, in order to portray the work as a dynamic object rather than archived in stasis on paper. Such methodologies and techniques are known under the umbrella term of the “digital humanities”. This research situates itself squarely between the musicological and the digital humanities, and endeavours to use the nascent digital humanities to research the ND repertory. The second research question is therefore: to what extent can the digital humanities inform, guide, extend upon, reform, and ultimately present, research of the ND repertory?

[14] Prime examples of this can be found in Tischler (1988) and van der Werf (1989).
[15] Rankin’s review of a modern edition of the collection of two voice substitute clausulae in **F** hypothesised that usage of that particular edition would require ‘an abundant supply of bookmarks’ (Rankin 1997, p.140).

1.2. Methodologies

As this dissertation is in part methodology focused, it is best here to summarise the ways in which the research questions outlined above will be tackled. This research is very much a part of the digital turn in the humanities. The ubiquity of computers in everyday life means that the function and utility of computers are fundamental in research and cannot be ignored: the use of computational methodologies can and does improve humanities research (Urberg 2017). In fact, this ubiquity can make the above point somewhat moot: it would now be almost unthinkable to complete humanities research without using a computer in some form. Smithies (2015, p.2) underscores the centrality of computing in modern humanities study by correctly asserting that to deny the use of computers in arts and humanities research 'is so far removed from common sense that it provides evidence of a distorting *mentalité*'. More than this however, the digital turn is mentioned here explicitly because the methodologies employed in this research are intensely computational, and at times completely new. This study is therefore not digital humanities research by convenience, but by choice.

The positivist methods advanced by early scholars in this field could also be regarded as computational, despite their flaws. Busse Berger (2005, p.10) has commented that Ludwig's work, from his earliest study calculating velocities of crusade-era troop movements, 'is concerned only with narrowly defined questions of fact that can be directly answered by reading the sources and he keeps speculation to an absolute minimum'. This at its heart is computer work, although the models used in the pre-computer age were small and their significance overstated. In this way, the modern digital humanities is similarly a turn back to that "objective", empirical mode of enquiry, hopefully a little wiser to the errors of the past, and a little more reticent to come to singular conclusions without knowledge of the bias of the data, research and researcher than some of the more inflexible research of the early twentieth century.

This is the same impetus that drives the digital humanities: 'getting things done' instead of endless speculation and theorising (Cecire 2011). The digital humanities is evidence- and results-driven. This research adopts those same ideals, albeit with a nuanced step against returning to black-and-white positivism. It also aims to "get things done" by using and, where necessary, implementing from scratch digital methodologies in order to find answers to the research questions. However, rather than arrogantly asserting that as a result this dissertation will unearth previously undiscovered "truths" of the ND repertory, this research takes a more careful scientific step in seeking not only to quantify its answers, but also the level of significance. It must not be forgotten that although I will be analysing the ND repertory using machines, the sources of the repertory were created by humans, and therefore the mode of research must be humanistic at its core. It will not suffice to interrogate the repertory by the means of binary distinctions as it is unlikely that the music can be classified into an unambiguous and complete taxonomy.

Rather than coming down on one side or another, this research strives to create evidence that demonstrates the shades of grey in between binary distinctions. This will be done by answering the research questions in terms that demonstrate the degrees of musical reuse rather than categorical decisions of concordance. In the same way, digital humanities researchers are often too focused on the power of the computer and allow their research questions to be aligned to what they know the computer can do (Berry and Fagerjord 2017, p.55). Rather, the computer is a methodological tool that can aid the understanding of questions already formulated: instead of bending the questions to fit the computer, the questions should be thought of first, and then the computer considered as an aid, even if it requires completely new methodologies to gain these understandings, or the dismissal of the computer entirely.

Often, the first problem for a digital humanities project, particularly one that deals with musical data, is in getting that data into the computer to begin with. The ND repertory resists simple encoding when we use current tools that are focused on common Western music notation (CWMN). In fact, there are only two studies identified that have sought to encode ND notation in its original notational context (that is, without transcribing to CWMN first): Erickson (1970) who used IBM punch cards to align organum purum in a tabular format, and Stutter (2020) where I used a hierarchical encoding format of my own design. As will be explained in "Chapter 3: Data models", neither tabular nor hierarchical approaches are sufficient for ND notation. However, one starting point other than this logical domain is the graphical domain of the central sources of the repertory (**F**, $\mathbf{W_1}$ and $\mathbf{W_2}$), now published as high-resolution facsimile images online by their host institutions with permissive licences (Biblioteca Medicea Laurenziana n.d.) [16]

Burgoyne et al (2016) list the four types of input data that a musical source for digital humanities research could take: images, symbolic data, audio, and metadata. The facsimile images of **F**, $\mathbf{W_1}$ and $\mathbf{W_2}$ are of the first kind. However, in order to perform meaningful analysis on the semantic content contained within, the graphical domain of the facsimiles must first be converted into symbolic data and stored in a music information retrieval (MIR) system alongside some added metadata. This is the first methodological hurdle: the conversion of images into symbolic data, and solutions to this have previously been either manual entry (keyboarding) or an optical music recognition (OMR) pipeline. The former is the simplest to implement but the most time consuming due to the manual labour involved, and the latter an enormously-complex, relatively

[16] **F** is public domain due to its age and facsimiles are hosted on IMSLP (International Music Score Library Project 2013), $\mathbf{W_1}$ and $\mathbf{W_2}$ are explicitly released by the Herzog August Bibliothek as Creative Commons (CC BY-SA) (Wolfenbütteler Digitale Bibliothek 2013).

new technology typically off-limits to all but the largest, most well-funded MIR projects.

The second methodological problem is that of analysis. Reflecting the above-stated desire to bend the technology to fit the questions rather than the other way around, the question of how to take a broad and wide ranging repertory such as ND polyphony and successfully segment its music into discrete units that reflect musical reuse is wholly novel. Furthermore, as ND notation and its related problems are almost entirely unsupported within current music analysis frameworks, the analysis support does not exist to load a ND notation file (whatever that may look like), unless it is first converted (i.e. transcribed) to CWMN. It is possible also that ND notation may use completely different metaphors in its notational style to CWMN and may not be adequately converted. Current analysis frameworks work on the implicit assumption that pieces of music are unrelated, and few tools allow for the analysis of two (let alone an entire repertory of) pieces. If we are to investigate questions relating to the segmentation of settings into clausulae, musical reuse and the interrelationships between settings, then an analysis framework that supports these high-level tasks must be used. The implementation of this research's methodology will provide such a framework.

1.3. Interdisciplinary research

This topic consists of a combination of the subjects of musicology, medieval source studies, and the digital humanities, and so this research is inherently interdisciplinary. One of the main hurdles that this study must surmount exists in the translation between the source materials (in the real world) and how they are modelled (in the computer). Put another way, one of the central issues in this research is one of data modelling. This includes creating a conceptual model that represents data in the computer, evaluating the efficacy of this model against real-world data, and determining that the model is useful to reason against the source material. This also requires discussion of how to then interrogate that model in order to answer the research questions (Ciula and Marras 2018).

This intersection also creates a difficulty with language for this dissertation: not only is it difficult to translate concepts between medieval musicology and the digital methodologies that I intend to use (which could be framed as a data modelling problem), but the technical terminology of both worlds must co-exist within one thesis — and often within the same sentence — which could render prose only half-comprehensible to all but a small group that are well versed in both specialities. This dissertation intends to ameliorate this issue somewhat by separating these two concerns as far as is possible into their own chapters and providing general summaries of each, keeping the thrust of the writing readable by experts of both main concerns as well as a more general readership.

Nonetheless, the aim of this research is musicological: the research question that concerns musical reuse is musicological and the digital aspect is introduced purely to support that question, not supplant it. The way back to an informed answer will therefore be by retracing our steps: a digital answer that will then be interpreted into a musicological conclusion. In the spirit of the digital humanities, it is only the methodology that is digital, the outer sections are musicological although the research is drawn from empirical methods.

Furthermore, it is my opinion that the "digital" aspect of the digital humanities particularly excels in methodology, such as building and creating (Carter 2013, p.3). When digital humanities is focused on methodology, the digital does not separate itself from the humanities (which emphasises writing and critiquing) but is instead simply another tool for the job of humanities research. This study takes a digital and computational methodology to the problems set out. However, this should not allow the "humanities" aspect to slip into the undertheorised, echoing Ramsay's cry of the persecuted digital humanist, guilty of a 'willful exogamy […] we're not quoting the usual people when we speak' (Ramsay 2013, p.244). Instead, this interdisciplinary study strives to straddle the gap and bring together both the digital and the humanities aspects without sacrificing the quality of either. Therefore, the aim is to quote the right people but then also go and make the tool; build and create in order to write and critique.

1.4. Another look

As a first step, then, let us take a new and fresh look at the central sources of the repertory — **F**, $\mathbf{W_1}$, and $\mathbf{W_2}$ — so that they might be best interpreted for digital methodologies. The physical characteristics of these sources, such as their size, shape, ink, and materials have been well described elsewhere and it would be unwarranted to repeat this information here.[17] Rather than these original physical aspects of the sources, this study is interested in the facets of notation and layout, and the music that that notation endeavours to represent.

What follows, therefore, is another look at the layout, notation, contents, focus, and construction of the three central sources from a standpoint that attempts to avoid placing ND polyphony in a position that would invite comparison with CWMN, in order to more objectively ascertain exactly what is and what is not known about ND notation. As has famously been said by Cook (2008, p.65), 'we are all

[17] **F** (Biblioteca Medicea-Laurenziana, Florence. MS Pluteus 29.1) is described in Dittmer (1967) and Ricci (2002), and at <https://www.diamm.ac.uk/sources/924/#/>, $\mathbf{W_1}$ (Herzog August Bibliothek, Wolfenbüttel. Cod. Guelf. 628 Helmst.) in Roesner (1974) and at <https://www.diamm.ac.uk/sources/870/#/>, and $\mathbf{W_2}$ (Herzog August Bibliothek, Wolfenbüttel. Cod. Guelf. 1099 Helmst.) in Dittmer (1969) and at <https://www.diamm.ac.uk/sources/854/#/>.

ethnomusicologists now' and, far from being limited to the music of so-called "others", the methods of ethnomusicology are just as successful when applied to 'ethnomusicology at home' (Nooshin 2011, p.285).

This new look therefore aims to describe the features of ND notation from an outside perspective without recourse to what might be familiar to us via CWMN. Traditional musicology has been critiqued for assuming too much and observing too little, and the advantages of starting our survey free from CWMN is that we must take nothing for granted (Cook 2008, p.52). As a result, this perspective is invaluable for the digital methods employed in this research as computational models, too, must often start from zero and build their knowledge without any prior domain knowledge. Even if there were a digital model that could "understand music" in the same way that we do, that model would also be biased in similar ways to our own internal models. To build our models from scratch, we must also consider ND polyphony from scratch, and make explicit everything that we want our model to implicitly understand. Moreover, it is particularly fitting to approach the sources this way when early music is commonly conceived 'with a keen sense of its distance, its unfamiliarity, its otherness' (Tomlinson 1993, p.4), while at the same time inviting favourable comparisons with composers such as Bach and Mozart (Busse Berger 2005, p.41).

As a result, my aim in this introductory survey is not to completely describe the sources of ND notation, but to describe their features in the most basic terms without reliance on CWMN. This survey may therefore appear extremely rudimentary to those with pre-existing knowledge of the repertory, but the purpose of these descriptions is to rely only on the sources as presented, accepting what it is we do not know, and not to consider ND notation as some anachronistic early form of CWMN. Moreover, such a fundamental a priori survey of these sources will become indispensable later on, due to the digital methodologies I will employ to extract, organise, store, and analyse this notation, as these will require exact data models for how the notation operates.

Some of the best pedagogical resources for gaining a knowledge of the operation of ND notation today are still Apel (1941, pp.215–281), Parrish (1957, pp.73–107), and Hoppin (1978, pp.215–255). Although these surveys are useful starting points for a student of ND notation, their aims are focused mostly on transcription, for example the conversion of modal rhythm into modern prescriptive rhythm. They will therefore not be that useful for this study as such practical transcription advice does not make clear the difference between: a) what can be extracted from the sources; b) what has been inferred from the writings of theorists; c) what is considered axiomatic to the study (such as a knowledge of modern notation to apply retroactively to medieval music); and d) their own opinions for transcription. These transcription guides are personal amalgamations of transcription skill rather than exact details of the operation of ND notation,

W_2 f.39v. Out-of-order systems

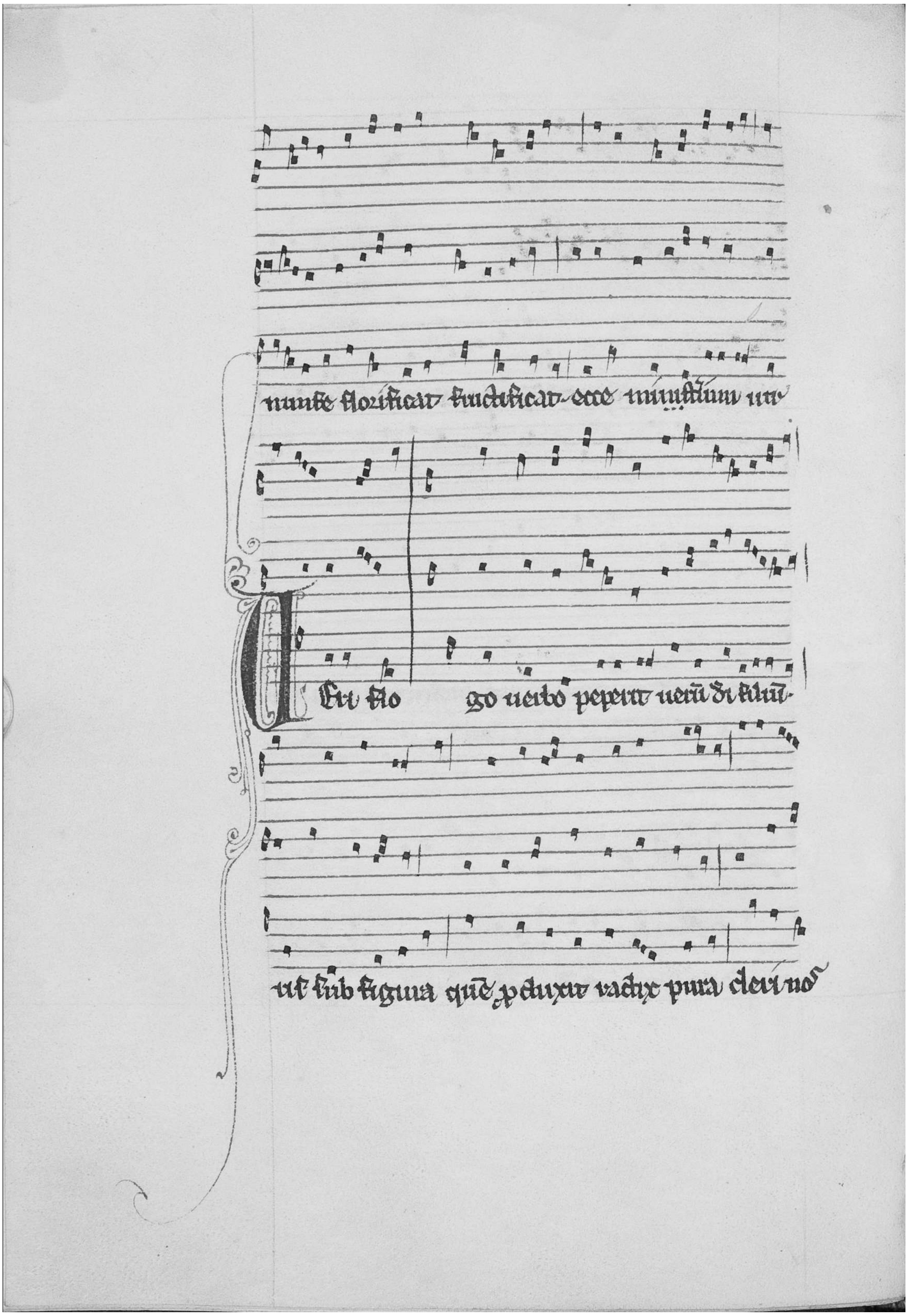

Fig. 1.2

leaning heavily on the idea of “modern equivalents”. For example, the introduction of rhythmic modes in all three cited here make short shrift of ligature patterning, diving directly into how those modes are best represented ‘in modern notes’ (Apel 1941, p.220; . The rejection of transcription for the purpose of this study is based essentially upon the same criticisms that are levelled at transcriptions of non-Western musics.[18]

On the flip side of this are histories and surveys of thirteenth-century notation, such as Hiley and Payne (2001), Busse Berger (2002a), Pesce (2011), Desmond (2018), and Earp (2018). These surveys are much more useful for this study, as they are based not upon a modern reading of theory and notation with the aim of translation, but a more in depth discussion of what that theory and notation means. Their findings have been used, in part, to construct this survey. However, it is important to understand that these writings form parts of broader contexts, such as histories of notation, and therefore necessarily have a retrospective view. For example, the subset of square notation used within ND notation is not explained in these surveys, as ND polyphony is conceptualised simply as a rhythmic and polyphonic form of square notation.[19] It is therefore understandable that these writings omit the fundamentals of the operation of ND notation — those fundamentals had perhaps been explained in earlier chapters of the comprised volumes — but what I wish to achieve here is, for the first time, to combine the fundamentals of what we know about ND notation outside of the context of Western notation in which it is typically understood. This is important, as if I am to create digital models that understand ND notation in order to extract patterns of musical reuse, then I must first describe it in its most basic form. Therefore, rather than considering ND notation historically as a proto-mensural notation, the goal here is simply a *prima facie* description of the source notation with minimal interpretation.

1.4.1. Layout and notation

Each source is written within the confines of a small writing block (**F** approx. 87×147mm, $\mathbf{W_1}$ approx. 91×160mm, $\mathbf{W_2}$ approx. 77×116mm), and within the page the notation is laid out in score format with a number of stacked staves each corresponding to a voice, except for some motets in $\mathbf{W_2}$ and **F** which have a tenor written separately to the other voices, often at the end or to the side, or exceptions where three-voice music is compressed onto two staves (see Fig. 1.7). Each line of music (which for clarity will anachronistically be termed a "system") is separated from the next by the combination of three factors which must be understood together: the existence of text, a slightly larger spacing, or the lowest voice moving slower than the upper voices (the tenor). These systems are always complete: never does the upper voice of a two-voice system appear on the lowest stave of a folio and the lower voice on the first stave of the next folio or its verso. Music moves from the top of the page to the bottom in system order. Within systems, each stave is simultaneous with the staves above and below, from

[18] Such criticisms are summarised in Nettl (2005, p.74–91).

[19] For example: 'the music is notated in the square notation of plainchant' (Hiley and Payne 2001, p.119), and 'the notation of our earliest sources of Parisian organa dupla ingeniously borrowed figures from contemporaneous plainchant notation' (Earp 2018, p.674).

W_1 f.32v, II-IV. Margin insertions

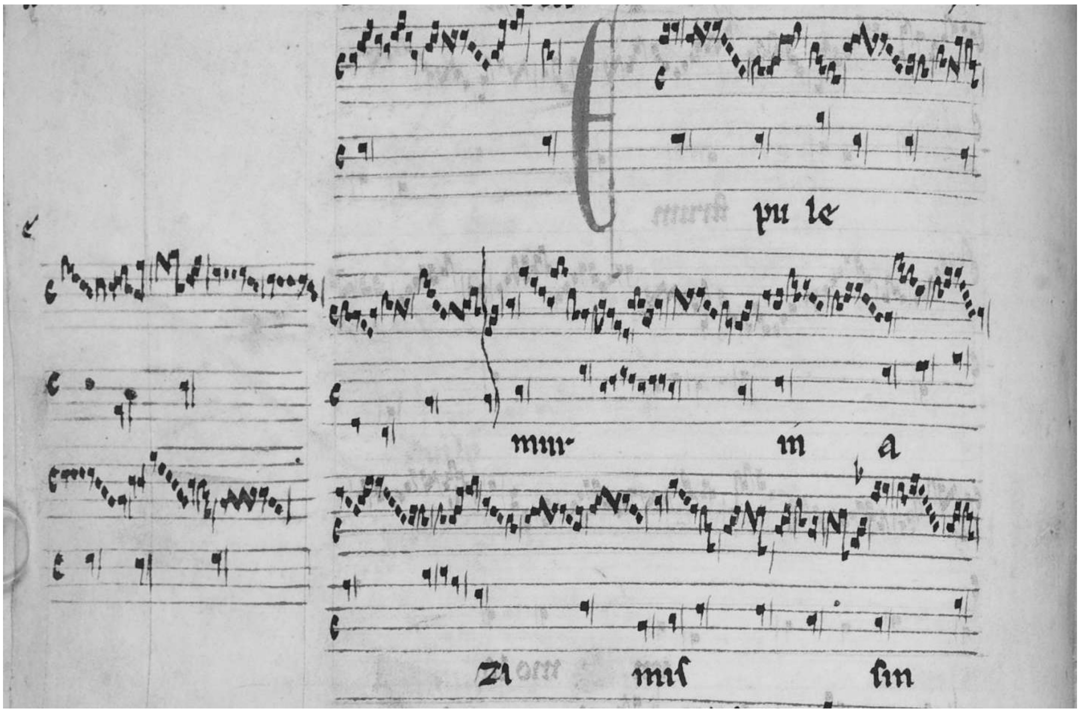

Fig. 1.3

left to right. There are two exceptions to this order.

The first is when an illuminated first letter must occur at the beginning of a line, but the previous setting of music has a passage still to be written. The overflow is written to the right of the first letter, effectively splitting the system in two. In Fig. 1.2, the text runs top-to-bottom, left-to-right thus:

> *nuntie florificat fructificat ecce ministrium uir-*
> *ueri flo- -go uerbo peperit uerum dei filium.*
> *-ris sub figura quem produxit radix pura cleri nos*

However, it should actually read:

> *nuntie florificat fructificat ecce ministrium uir-*
> *-go uerbo peperit uerum dei filium. Ueri flo-*
> *-ris sub figura quem produxit radix pura cleri nos*

It is clear to see that the vertical line in II splits the system in two, the right hand side completing the transmission of *Isayas cecinit*, the left beginning *Ueri floris sub figura*.

The second is when the scribe has made an error, missing out a passage of polyphony from one or more voices. The scribe adds extra system(s) in the margin which are typically read from top-to-bottom and inserted at a mark in the errant system. In Fig. 1.3, a crooked vertical line in III indicates that there is something missing, and

W_1 f.10v, II-IV. *Natiuitas gloriose* with alternate text *[Opti]mam partem*

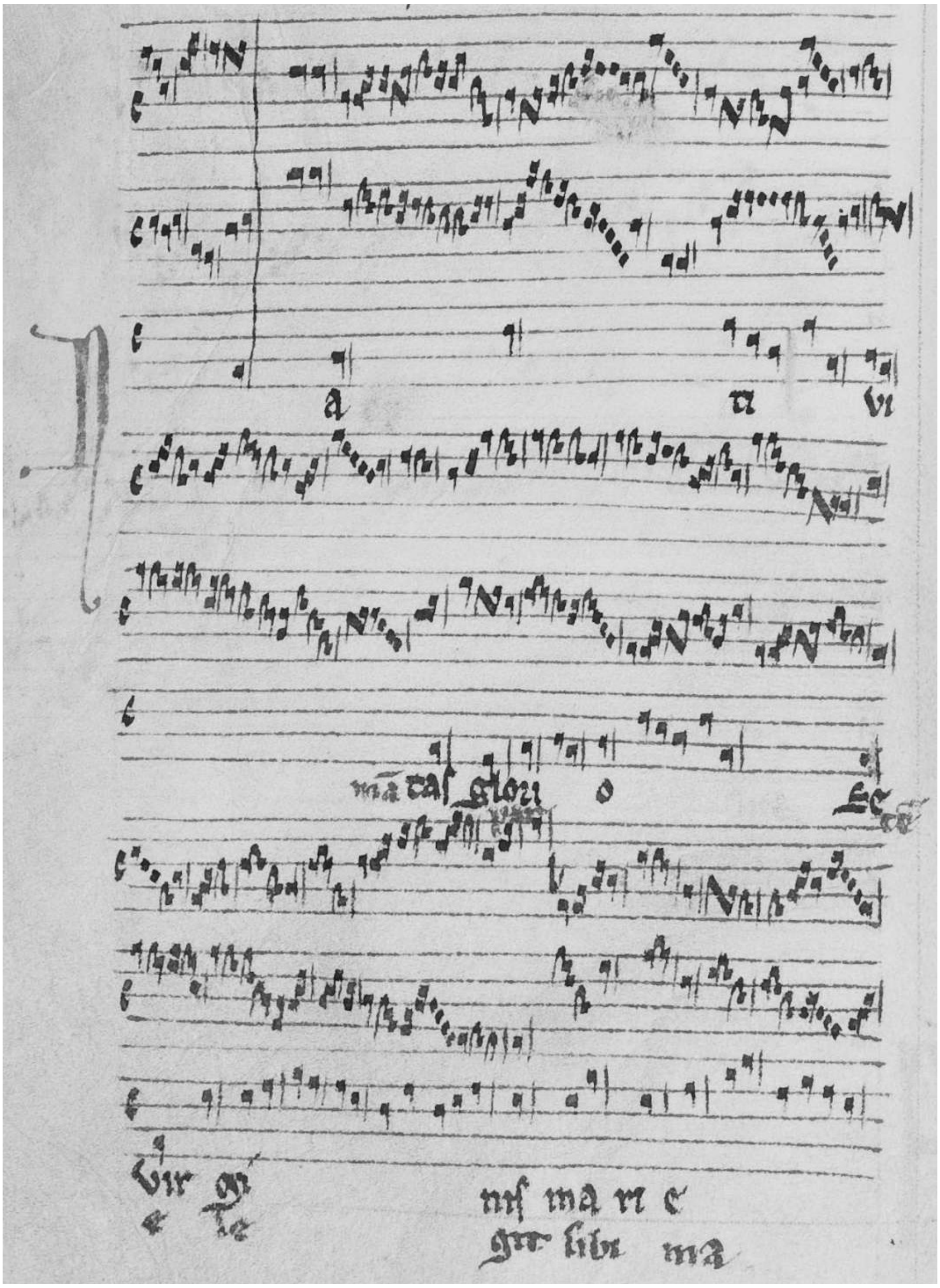

Fig. 1.4

the two margin systems are inserted at this point.

Within a system, all the musical events occur within the same time frame. For example, a long note in a lower voice will last long enough for all the notes written above, and two notes in an upper voice will move at twice the speed of one note in the lower voice (although not necessarily with equal duration). This can be confirmed by collating music that occurs in multiple manuscripts with different system breaks. The two exceptions to this are: a) when a long note is continued over a line break by a large space at the beginning of the next stave (this typically happens only in tenor voices); and b) when a setting of music begins halfway through a system. Logically however, this last case should be conceived of as two separate systems, one for each setting.

Generally, each system and setting begins with a clef. Although there are a few notable exceptions, these can invariably be chalked up to errant scribal omissions. There are three kinds of clef in use: C clefs of all kinds (most commonly C3 and C4) used in both tenor and organal voices, F clefs of all kinds (typically between F3 and F5) used only in tenor voices, and a single instance of a D clef in $\mathbf{W}_1$ f.69r, III. The line that the clef sits on does not indicate the absolute disposition of the voice but rather the relative pitch of that stave or passage, often designed in such a way as to avoid ledger lines. As a result, it is not uncommon for a high-pitched passage to be written in a high clef and a subsequent low-pitched passage in a lower clef, or vice versa. B-flats, E-flats and F-sharps that form part of the mode can be expected to occur at the start of the system, but *ficta* are written just before — although not always immediately prior to — the note in question. However, it can be expected that a sharp or flat then followed by its opposite will be cancelled. A system break cancels all *ficta*.

Text occurs at the bottom of a system, with the exception of extra stanzas for conductus which usually occur at the end of a setting. Occasionally there are multiple lines of text under a system, indicating an alternate text setting (see Fig. 1.4).

Dark square or rhombic note heads indicate a pitch by their vertical location on the staff lines in context of the clef as well as any previous flats or sharps. When a desired pitch falls outside of the scored staff lines, extra staff lines are temporarily added to accommodate these higher or lower notes. It would therefore be unfair to term them ledger lines as they are not intermittent, one for each note, but are drawn parallel to the staff for the duration of the passage. More commonly, notes simply move outside the bounds of their stave onto the stave above or below; staves do not necessarily "belong" to a voice. Pitch can also be inferred by the direction of *plicae*, which are upward- or downward-pointing stems at the end of a note. This is with the exception of single *puncta* or *virgae* which may already have a stem. In the case of downward *plicae*, a stem is also added to the left side of the note head to indicate that the right stem is a *plica*. For upward *plicae* either the note head: a) has no stem, save that of the *plica*; b) a stem is added to the lower right as well, giving a joined stem that runs all the way down the right hand side of the note; c) a stem is added to the left of the note, similar to that of downward *plicae*; d) is equivalent to a flipped downward *plica*, with two upwards stems on left and right (see Fig. 1.5).

Notes are often ligated into groups. Some of these ligations, particularly those with

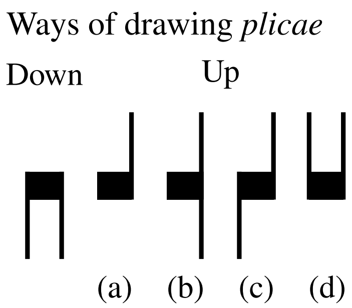

Fig. 1.5

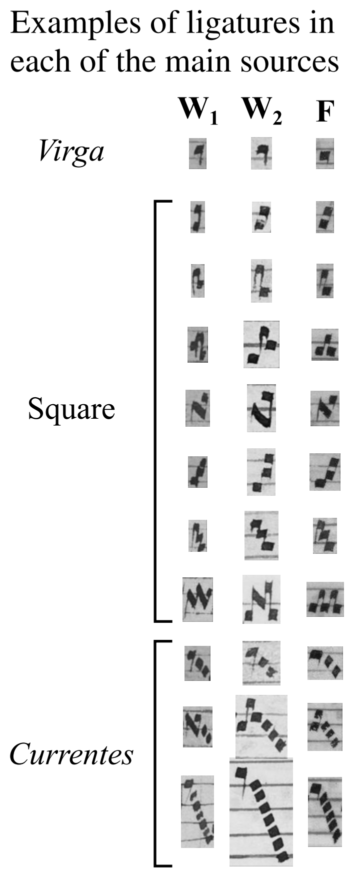

Fig. 1.6

the most regular patterns, inform a rhythmic interpretation of the notation (*sine littera* notation). However, although it is generally understood that the groupings of ligatures may manifest a rhythmic mode that can then be applied to the music, exactly to which settings of music these rhythmic modes are applied and how they are to be applied in the case of *fractio modi* is unknown. This is especially the case in purum, where scholars have had numerous disagreements on if and how rhythm should be applied to these passages.[20] It is generally agreed that passages strictly adhering to a regular patterning of ligatures are rhythmical, and it makes sense that three- and four-voice music — which must require exact timing in performance — is typically written this way. Similarly little understood is the rhythm of *cum littera* notation in motet and conductus, where the poetic meter of the text controls the rhythm of the music, and ligatures are untyped. In this notation, a ligature indicates only how many notes fit within that longa or breve. Once again however, the rhythm of these semibreves in ligature is unknown.

There are two forms of ligature: square and *currentes* (see Fig. 1.6). Square ligatures are formed by square or *flexa* pen strokes joined by straight lines at right angles,

[20]For a summary of these issues, see Everist (2003, pp.lxxi–lxxiv) who, after Yudkin (1983), attributes this disagreement to a faulty edition of Garlandia, and in his edition occasionally provides ‘rhythmic suggestions above the stave’ for those purum passages whose notation implies a particular rhythmic interpretation (Everist 2003, p.lxxxv).

and *currentes* by any number of rhombic note heads, usually beginning with a single *virga*. *Currentes* are employed typically to encode descending scales of three or more notes, but can be observed to move upwards or in both directions, as well as moving by leap rather than step. *Plicae* can occur at the end of square ligatures, but have also been observed in rare cases in the middle of a long ligature. The forms of ligatures are generally the same as ligatures found in common diastematic chant material of the period. It is not possible to ligate two adjacent notes of the same pitch, but repeated notes must often be read as a single ligature. Ligatures are contextually parsed, that is, notes do not need to be physically touching on the page to be considered a ligature — especially in the case of *currentes* — but are considered ligated by their proximity to other notes. This can be difficult to read in particularly cramped sections of notation as there often can be no discernible gap between adjacent ligatures.[21] In some hands, the first note of a square ligature may have a small, downwards-pointing stem on its right hand side which may indicate that that first note in the ligature is to be interpreted as a longa, but these extra markups are inconsistently applied and cannot be relied upon for rhythmic interpretation. A ligature may move from square to *currentes* notation, especially when that ligature ends with a rapidly descending scale of notes, but it is once again unclear whether this should be considered a new ligature or an extension of the previous. The question remains whether *currentes* must begin with a *virga* to be considered a new ligature, or whether *currentes* can be preceded by a number of notes in square ligature. This issue can be avoided by considering only the rhombs themselves as part of the *currentes*.

Sets of ligatures are found in groups commonly called *ordines* where a small vertical line (a *divisione*) indicates their division. *Divisiones* can be found at any point in simultaneous voices: together as a rest and gathering-together point, and separate in a hocket style. However, the only *divisiones* that can be guaranteed to occur simultaneously are those which are drawn through multiple staves: all other *divisiones* are based on context, concord with other voices, and rhythm.[22] At the same time, some *divisiones* appear to reflect not rests or pauses but are simply markers of section and syllable changes.

1.4.2. Contents

It is a point of contention whether **F**, $\mathbf{W_1}$ and $\mathbf{W_2}$ are attempts to create copies of the same book, or multiple books with some overlapping content. Central to the study of

[21] For example, compare Payne's reading of the *Gloria* of *Constantes estote* (O1) from **F**, f.65v (Payne 2021–, fasc.3, p.3) to Everist's reading of the same (Everist 2003, p.5). Whereas Everist reads the section "-ri-" as ligatures 3+4+3+4+4+2, Payne reads 7+7+4+2, presumably interpreting the close notation as longer, combined ligatures.

[22] This contextual alignment at times completely alters the interpretation. See Karp (1966, pp.350–352).

ND polyphony are the writings of a late thirteenth-century English author known in modern study simply as "Anonymous IV" (Wegman 2015). Anonymous IV introduces many of the terms that we now use in the study of ND polyphony, including the term *Magnus liber organi* which can be translated to 'great book of organum' (Yudkin 1985, p.39). This alludes to the existence of a singular book, possibly copied multiple times as a source for all extant copies, but the central sources of **F**, $\mathbf{W_1}$, and $\mathbf{W_2}$ exhibit differences between their transmission of settings of music far beyond what would be expected for such a purely literate manuscript culture.[23] As a result, modern writers use the term *Magnus liber organi* variably, based on their personal interpretations of what Anonymous IV meant by *Magnus liber organi*. The term may refer, most narrowly, to only the "un-redacted" cycle of two-voice organa present as a central layer in all three sources.[24] It may also refer to the combined collection of organa and clausulae, perhaps with the largest source **F** as a central text.[25] Occasionally it may refer to the unified set of these two genres between the central sources, including *unicae*.[26] It may also be a broader definition of thirteenth-century polyphony, including conductus,[27] or even an expression of any polyphony without limitations of genre.[28]

It is difficult therefore, to use the term *Magnus liber organi* to succinctly pin down any part of this repertory. Insofar as the term referring to a book — written down or otherwise — common between these central sources, it should suffice to say that no contiguous section of any source is replicated elsewhere. Although some polyphony may be apparently copied verbatim between one source and another, very rarely do the same settings of music appear sequentially in two sources. Exceptions to this, such as the four-voice *Sederunt principes* following the four-voice *Uiderunt omnes* are most likely simply a common sense of organisation in rite and polyphony, as well as convenience. The scribes of each central source cannot have been following some common organisational schema such as a single exemplar, but organised their source as they saw fit and as materials became available to them. In order to draw evidence to this, let us examine the contents of each source in detail. I will not delve too deeply into the codicology of

[23] See the discussion of alternate transmissions of ostensibly the "same" music in Roesner (2001, pp.242–246).

[24] 'The *Magnus liber organi de gradali et antiphonario* is a collection of responsorial chants of the Office and the Mass in the form of organum duplum' (Waite 1961, p.275).

[25] 'The existing sources represent different selections from a common repertory pool, which existed in separate copies of the several organa and their sections and often in several variants' (Tischler 1984, p.168).

[26] 'Three further, different groups of pieces expand this basic corpus, namely those common to F and W_2, which also contains no *unica*, those common to W_1 and F mentioned above, and finally the large number of *unica* in F' (Husmann and Reaney 1963, p.314).

[27] 'The "great book" that Leoninus is said to have "made", the so-called *magnus liber organi*, is, of course, the great collection of liturgical polyphony and monophonic conductus prepared for Notre-Dame' (Roesner 2001, p.231).

[28] 'The *magnus liber* is generally regarded by present-day scholars as consisting of *organum duplum* and clausulae. However, it seems likely […] that the *liber* was not limited to a single genre […] but rather included compositions in all genres cultivated by the musicians of Paris — organum, conductus, and the motet' (Roesner 1993, p.lix).

each source (gatherings etc.) except when it pertains directly to the content (such as missing folios etc.).

F is the largest of the three sources studied, consisting of a foliation extending up to 476. It is also by far the most well organised of the sources, likely as a result of it being written by a single hand. As can be seen in Tbl. 1.1, **F** proceeds generally from organa to conductus to motet, and from four-voice to three-, two- and one-voice music.[29] The vast majority of the substitute clausulae are collected together in fascicle V. The two exceptions to this schema are firstly Fascicle I which gathers together all the four-voice music of the repertory at once, rather than dispersing its genres across the other sections of the MS. Secondly is the case of the nine substitute clausulae from f.10v to f.13v that immediately follow: the two-voice *Tamquam* as well as the three-voice *Go*, *Flos filius eius*, *Domine*, *Eius*, *Diffusa est*, *Descendit de celis*, *Gloria*, and *In seculum*.

These small exceptions notwithstanding, and accepting the unknown contents of **F**'s missing folios, it is clear to see from the well established order and contents of **F** that the scribe of this source had all his materials available to him — or, at the very least, a good idea of how much music he could procure and when — before he began writing. Unlike $\mathbf{W_1}$ which often has to fit music onto folios of the wrong size or layout, resulting in cramped or messy notation, **F**'s good organisation results in neither of these issues, and can afford many extra folios of ruled staves to add future music. This is especially clear when we see that each fascicle is — with the exception of Fascicle I — dedicated to a particular genre of music. **F**'s focus is that of conservation and monumentalisation, and organises its contents as a presentation of the repertory in a directory format. The scribe of **F** was not copying sets of transient music that passed by him in order of receipt as a working copy to use later, but was attempting to preserve music from sources (literate or otherwise) that were easily accessible and readily available to him and could be copied in any order.

When it comes to organisation therefore, **F** is a model source, as it organises its materials almost entirely from the sacred to the secular and from the many-voiced to the few-voiced. Indeed, it only breaks the organisational cycle once when the liturgical calendar is reset for the two-voice organa cycle beginning Fascicle IV.

At the other end of the spectrum to **F** is $\mathbf{W_1}$ which, though it can be seen to attempt some form of content organisation, must frequently break its schema in order to add more music. As with **F**, $\mathbf{W_1}$ begins with those four-voice pieces, although the first two folios are missing, presumably taking *Uiderunt omnes* and a large part of *Notum fecit* with them. Tbl. 1.2 shows that although $\mathbf{W_1}$ begins with the four-voice repertory and

[29] The tables in this section are mostly précis, with amendments based on study of facsimile images, from the dense prose of Ludwig (1910) and the tables in Smith (1964).

W_1 f.63r. Three-voice music on two-voice staves

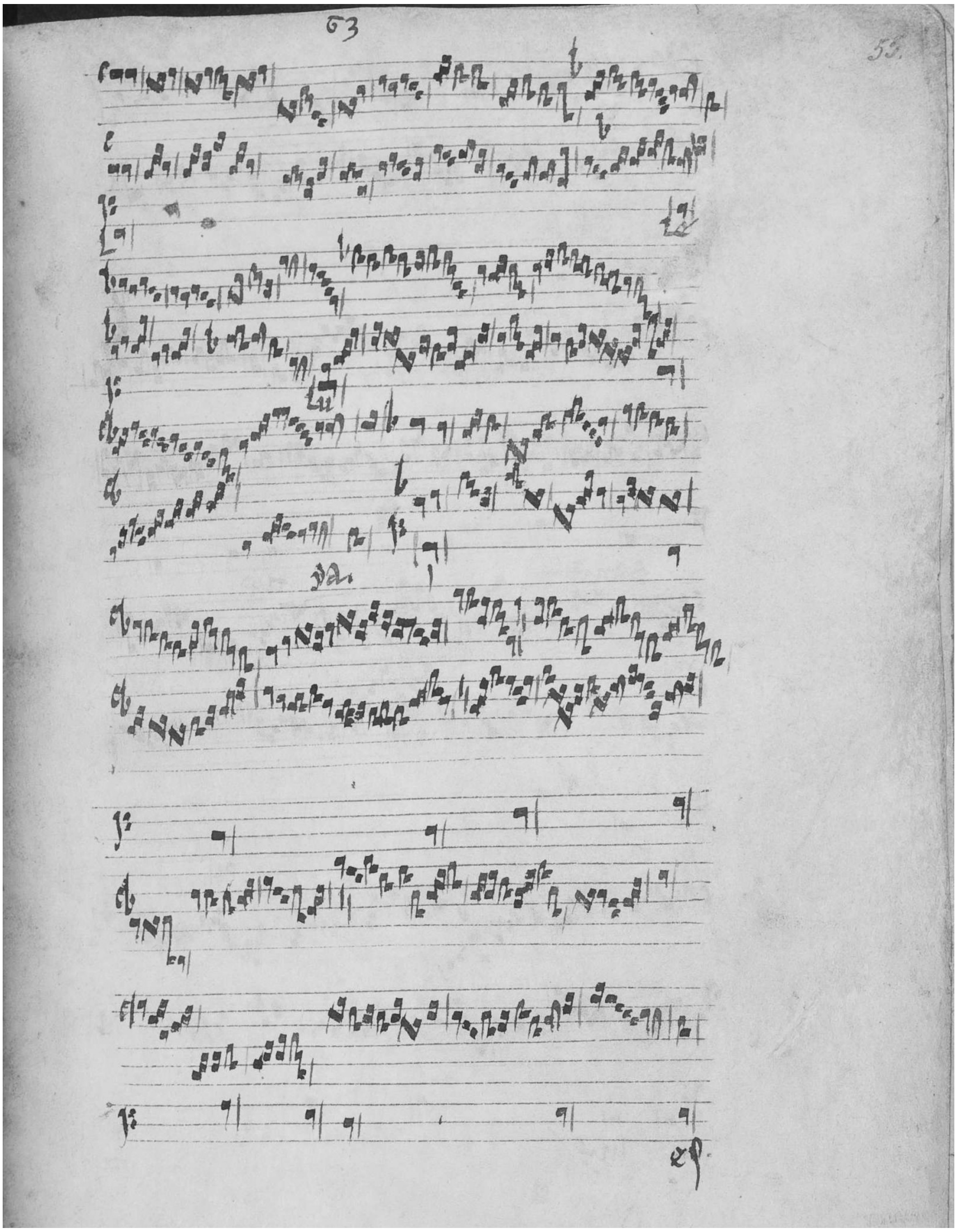

Fig. 1.7

some three-voice music as well, and successfully begins the second cycle of organa at f.25r, f.63r heralds what must have been an entirely unplanned schema shift in a new gathering, as a set of three-voice organa is transmitted, three voices squeezed onto two staves (see Fig. 1.7). From there the schema begins to break down more fully, as organa, conductus, and Mass Ordinary tropes are inserted, seemingly without much high level organisation. It is from here, too, that the scribal hands change more frequently and issues of cramped and non-aligned spacing become more prevalent. We can thus consider the first six fascicles of **W$_1$** (the first eight gatherings) as a more unified repertory layer that was perhaps organised as one, followed by further layers that were either added to the MS at a later date or were bound together with it.

The number of rows of this table indicates that **W$_1$** is a much less well organised source than **F**. Despite **F** being a much larger collection than **W$_1$**, the latter creates more rows in the table due to the frequent changing of genre and number of voice parts. After a fairly well organised first six fascicles, Fascicle VII begins similarly with three-part organa and Fascicle VIII with two-part organa, but this schema is then abandoned and the set mixes three-voice organa, conductus and Mass tropes together. Fascicle IX appears to have been conceived as a conductus set — a large part of its opening folios is devoted to three-part conductus with two-part *secunda pars* — but it then devolves once again into a broad mix of genres. Fascicles X and XI are a similar story, but it is interesting to note that by f.106r all organa have been exhausted. Despite the apparent disorganisation, then, there is a sense in **W$_1$** of organa being transmitted before the "lesser" genres of conductus or Mass tropes, and **W$_1$**'s only set of substitute clausulae (f.49r–62r) following directly from its set of two-voice organa, just as in **F** f.147r.

Nonetheless, clearly the scribes of **W$_1$** did not have the same luxury as the scribe of **F**'s access to materials, perhaps due to **W$_1$**'s provenance being peripheral to the repertorial centre in Paris (Everist 1990). As a result, **W$_1$** appears to be organised from an initial set of materials, with new materials copied into whatever fascicle was incomplete at the time, as they became available.

Between these two organisational extremes lies **W$_2$**, which is not as well organised as **F** nor as apparently disordered as **W$_1$** (see Tbl. 1.3). Like **F** and **W$_1$**, **W$_2$** begins with a fascicle of four-voice music followed by three-voice music. However, **W$_2$** appears to be more focused upon transmitting motet than organa and dispenses with its organa within the first five fascicles. Nor does **W$_2$** contain substitute clausulae, transmitting only its version of the canonical organa settings. Fascicle VI is dedicated to two-part conductus and VII to mostly three-part Latin motets and conductus. However, Fascicles VIII through X almost exclusively contain motets. Fascicle IX is an unordered set of French three-voice motets but VIII and X are each organised into three sets of alphabetical order.

Such alphabetical order is in fact a rather modern ordering technique for the period, indicating perhaps that **W$_2$** is the youngest or most progressive of the sources. Fascicle VIII first contains two alphabetically-ordered series, mostly of two-part motets. This latter series contains two exceptions to its organisation: firstly the one-voice conductus *Beata viscera* on f.156v which appears to have been errantly inserted into the alphabet of this motet cycle, and secondly a single three-voice motet: *Mors morsu nata — Mors que stimulo — Mors* on f.164v–165v. The final alphabetical series of this fascicle is interrupted by one more three-voice motet: *Salue salus hominum — O radians stella — Hostium*. There then follows seven more two-voice motets that fall outside of an alphabetical order.

Fascicle X is dedicated entirely to two-part motets in this same order, but the final alphabet — either by design, error or non-completion — only catalogues motets from A to D. There then follows three more un-alphabetised motets. However, an alternate possibility could be that this final alphabet is not partial but simply quite short, and no motets of letter E–O were available: the final three “un-alphabetised” motets beginning with letters P, S and T respectively, forming one discontinuous alphabet.

$\mathbf{W_2}$ therefore demonstrates an interesting intermediary between the fully-planned-out **F** and the more mixed collection of $\mathbf{W_1}$. Each fascicle has been planned in some way but more often than not with some exceptional elements. The first six fascicles and the opening of the seventh proceed in the same hierarchy as **F**, indicating a single layer. Fascicles VIII–X are an independent unit focused on motet, but it seems as though these final three fascicles were written in at least eight or nine independent stages, one for each alphabet or un-alphabetised section.

1.4.3. Construction and musical reuse

Of the genres of thirteenth-century polyphony, it is generally believed that motet is the latest genre in this collection, given its extensive borrowing from organa and the motet’s dominance over the relative paucity of organa and conductus in later fourteenth-century collections (Bradley et al 2019). Conductus are organised quite simply in multiple stanzas, and each voice moves mostly note-against-note in *cum littera* notation. The construction of conductus mirrors that of the poetry it sets, but there is some scant evidence of borrowings between conductus and clausulae, and intertextuality within conductus (Bukofzer 1953).

Organa have a much more complex construction, owing to their multiple sub-genres and the almost completely inscrutable network of borrowing in substitute clausulae and beyond. Tenor voices are typically the lowest part and either hold a long note (commonly called purum), or move in modal rhythm (discant). There is also copula, which is an ad hoc mixture of the two styles. In the organal voices, it is not known to what extent the movement is subject to modal rhythm. It is clear in discant as well as in some copula that the ligature patterns imply a particular modal rhythm, but regular patterning in purum is not always present.

Older editions, such as those by Waite (1954) and Tischler (1988) used modal rhythm throughout, as they interpreted unpatterned purum in modal rhythm, in a similar way to discant (Waite 1954, p.119). However, especially since Yudkin’s article (1983) advocating for free rhythm in purum based on a re-reading of the theorists, transcriptions generally interpret purum as free, except in those cases where patterns are clearly apparent in the duplum. However, where this line between “clear” versus “unclear” patterning sits depends on the editor. For example, it has already been mentioned (on p.24) that Everist’s (2003) transcriptions of the duplum of **F** use suggested rhythms above the stave for these moments of clearly-patterned purum. His transcription of *Iudea et Iherusalem. Constantes estote* (O1) demonstrates this (Everist 2003, pp.1–6), where the opening is transcribed arhythmically, but the beginning of the syllable “-stan-” has a suggested rhythm in mode 1, meaning that Everist has believed the notation of this passage to be patterned. Comparing this to Payne’s (2021–, fasc.3, pp.1–4) transcription of the same, we can see that he agrees on the patterning of “-stan-”, but opens with rhythm where Everist has none. This is generally the case between these two editions, where Payne’s threshold for what constitutes patterned notation is lower than Everist’s.

Substitute clausulae have long been known as substitutes for passages of organa, and Ludwig (1910) diligently split those canonical organa into their clausulae components. There was some dissatisfaction at the time with the way Ludwig arranged clausulae in order to reflect what we now know was an invented chronology, most notably from Handschin (Busse Berger 2005, p.34), but it was not until Smith (1964, pp.7–12) that Ludwig’s taxonomy was critically dissected. Smith’s main criticism was that Ludwig’s selection was geared towards discovering the origins of motet, not of organa and clausulae themselves, and so ‘his treatment must also be viewed as unsystematic and very incomplete’ (Smith 1964, p.10). Smith instead selected clausulae from all settings of organa, canonical or otherwise, without Ludwig’s concern for which end up in motet and which do not, and collated his findings into a complete set of tables and diagrams that combine settings and isolated clausulae together by chant and liturgy, culminating in repertorial diagrams of settings that reconstruct organa, not as a collection of individual and static “compositions”, but as the sum of their component sections. No organum is excluded. Smith demonstrated that, without exception, organa can be viewed as a concatenation of passages: purum, discant, and substitute clausulae.

This culture of musical borrowing and reuse arguably culminates in the genre of the motet, which freely mixes original material with the music of chant, organa and secular sources. The tenor of a motet is often taken from a chant source, and an ND discant clausula can be converted into motet material (Kidwell 1998). Bradley (2013) has also shown that the reverse can be true: some substitute clausulae may have began life as motet. It is yet unknown to what extent this borrowing was commonplace. The question of “What did organa and motet borrow from each other?” was asked and answered by Smith (1964). However, the questions “What exactly made a particular passage or idea suitable for borrowing?” and “What other kinds of musical reuse existed outside of the clausula?" remain unanswered. Bradley (2019) has made tentative progress in determining the conditions necessary for chant to become motet tenors, but the conditions for other elements of motet — or indeed organa, substitute clausulae, or conductus — to be used as material for musical reuse are as yet undiscovered.

1.5. Structure of this dissertation

As can be seen, the concerns of musical reuse permeate the genres of thirteenth-century polyphony. Therefore, it is vital that any attempts to tackle the core questions of the repertory consider this tricky issue, not only because musical reuse appears to influence every aspect of ND polyphony, but also because study of the ND repertory has neglected this particular topic for so long. This dissertation attempts to begin to fill this lacuna by the application of new, digital methodologies to investigate it, not only through the known process of the substitute clausula but beyond, by investigating musical reuse at a broader level in indirect concordances and melodic formulae.

Concomitantly, this dissertation's structure reflects the separation of the musicological and digital concerns. Chapters with a broadly musicological focus discuss the musicological issues raised, and chapters with a methodological bent focus on the digital methods used. This is done: a) to avoid writing with an all-too-general tone throughout, systematically patronising experts of both disciplines; and b) to allow for different audiences on a per-chapter basis: digital chapters are not designed to be completely understood by musicologists (and vice versa), with the caveat that the main ideas of each section should be mutually intelligible. With that said, this dissertation can be divided into two strands: "discussing" and "doing". Using the terminology of the digital humanities, the humanities aspect (writing and critiquing) is considered in the first few chapters, followed by the "digital" (building and creating), before finally returning to the humanities questions that this dissertation will answer.

Therefore the first "discussing" section, following on from this introductory chapter, reviews two disparate ways in which the terms used in this research can be defined using current thinking. First, the issues surrounding musical reuse up until now will be introduced and discussed through a historiographical investigation into the term "clausula", and then a review of current digital music encoding methodologies. "Chapter 2: Historiography of musical reuse: the clausula and its context" looks in detail at how the term "clausula" has been co-opted and elevated in modern study as a byword for all musical reuse, *pars pro toto*, to the detriment of indirect concordances and melodic formulae. It surveys the literature of both clausulae and melodic formulae, examining how musicologists have defined and discussed musical reuse: the tools that they have had at their disposal and the evidence that they draw upon to make persuasive arguments. The question is asked whether single definitions of "clausula" and musical reuse can be drawn from the literature, and therefore whether such terms can truly be rigorously defined.

"Chapter 3: Data models" considers the data models and methodologies that are commonly used to encode, conceptualise and analyse music. It traces where these models came from, the thinking behind these encodings, the technological cultures that created these models and, as a result, the kinds of music for which they are particularly suited. Again, a follow up question is then asked: whether these technologies and models can reliably be used in their current operation to encode ND notation or analyse questions of musical reuse in the repertory and, if not, whether they can be moulded or bent to the task.

The second, "doing" part of this dissertation is primarily methodological. "Chapter 4: Requirements" summarises the conclusions reached in the first part of the dissertation and creates requirements for the methodologies created in the succeeding chapters: what needs creating to make a system for analysing musical reuse in ND polyphony, what must it do, what are the inputs/outputs and what is the purpose of this technology? There then follow two methodology-heavy chapters: "Chapter 5: Online" for the online component of the research (website for data input and presentation of output), and "Chapter 6: Offline" for the offline component (data analysis and statistical enquiry). Each chapter draws upon the conclusions and requirements of the "discussing" part such that the "doing" does not alter the parameters of what has already been decided upon.

"Chapter 5: Online" runs down the components, technologies available and technologies chosen for each part of the online system: database, server, client, and OMR. For each part of the system, the requirements are identified and a broad understanding of how those requirements are implemented is described. "Chapter 6: Offline" then completes a similar task for the data analysis, and outlines at a high level the features of the data analysis framework created. Statistical analysis is also discussed within this. There then follows "Chapter 7: Results and Analysis", which will evaluate the new methodologies created for the purposes of this study, analyse the results generated from the system, and the two threads of this dissertation are brought together. We then retrace our steps: the digital is used to create humanities (musicological) evidence. Finally, "Chapter 8: Conclusion" will sum up the results of this dissertation, the evidence, what can be learned from its conclusions, its limitations, and avenues for future research.

Contents of **F**

Folio	Genre	Voices
Fasc. I (1r–13v)		
1r	Organa	4
7v	Motet	4
10v	Clausulae	2
11r		3
Fasc. II (14r–64v)		
14r	Organa	3
48r–64v	***Missing folios***	
Fasc. III (65r–98v)		
65r	Organa	2
Fasc. IV (99r–146v)		
99r	Organa	2
Fasc. V (147r–200v)		
147r	Clausulae	2
185r–200v	***Missing folios***	
Fasc. VI (201r–262v)		
201r	Conductus	3
Fasc. VII (263r–380v)		
263r	Conductus	2
Fasc. VIII (381r–398v)		
381r	Motet	3
398v–?	***Missing folios***	
Fasc. IX (399r–414v)		
399r	Motet	2
414v–?	***Missing folios***	
Fasc. X (415r–462v)		
415r	Conductus	1
Fasc. XI (463r–476v)		
464r	Conductus	1

Tbl. 1.1

Contents of $\mathbf{W_1}$

Folio	Genre	Voices
Fasc. I (1r–8v)		
1r–2v	***Missing folios***	
3r	Organa	4
6v	Motet	4
7r–8v	***Missing folios***	
Fasc. II (9r–16v)		
9r	Organa	3
13r	Motet	4
15r	Conductus	3
Fasc. III (17r–24v)		
17r	Organa	2
24v	Tropes	2
Fasc. IV (25r–48v)		
25r	Organa	2
36r–37v	***Missing folios***	
38r	Organa	2
Fasc. V (49r–54v)		
49r	Clausulae	2
51r–52v	***Missing folios***	
53r	Clausulae	2
Fasc. VI (55r–62v)		
55r	Clausulae	2
62r	Conductus	2
Fasc. VII (63r–69v)		
63r	Organa	3
Fasc. VIII (70r–94v)		
70r	Conductus	2
72r		3
83r–84v	***Missing folios***	
85r	Conductus	3
85v	Organa	3
88r	Conductus	3
90r	Organa	3
91r	Clausulae	3

Folio	Genre	Voices
91v	Tropes	3
Fasc. IX (95r–176v)		
95r	Conductus	3–2
103v	Organa	2
104v	Tropes	2
105v	Organa	2
106r	Tropes	2
106v	Conductus	2
120v	Tropes	2
121r	Conductus	2
176v	Tropes	2
Fasc. X (177r–192v)		
177r–184v	***Missing folios***	
185r	Conductus	1
185v	Tropes	1
Fasc. XI (193r–214v)		
193r	Tropes	2

Tbl. 1.2

Contents of $\mathbf{W_2}$

Folio	Genre	Voices
Fasc. I (?–5v)		
Before 1r	***Missing folios***	
1r	Organa	4
5r	Clausulae	4
Fasc. II (6r–30v)		
6r	Organa	3
Fasc. III (31r–?)		
31r	Conductus	3
After 46v–?	***Missing folios***	
Fasc. IV (47r–62v)		
47r	Organa	2
Fasc. V (63r–91v)		
63r	Organa	2

Folio	Genre	Voices
Fasc. VI (92r–122v)		
92r	Conductus	2
Fasc. VII (123r–144v)		
123r	Motet	3
Between 133v–134r	***Missing folios***	
134r	Motet	3
138v	Conductus	3
140r	Motet	3
141r	Conductus	2
Fasc. VIII (145r–192v)		
a) Alphabet i (145r–155v)		
145r	Motet	2
b) Alphabet ii (155v–178r)		
155v	Motet	2
156v	Conductus	1
157r	Motet	2
164v		3
165v		2
c) Alphabet iii (178r–190r)		
178r	Motet	2
186r		3
186v		2
d) Un-alphabetised (190r–192v)		
190r	Motet	2
Fasc. IX (193r–215v)		
193r	Motet	3
Fasc. X (216r–253v)		
a) Alphabet i (216r–222r)		
216r	Motet	2
b) Alphabet ii (222r–248r)		
222r	Motet	2
c) Alphabet iii (248v–252r) (Partial A–D)		
248v	Motet	2
d) Un-alphabetised (252r–253v)		
252r	Motet	2

Chapter 2

Historiography of musical reuse: the clausula and its context

In a curious footnote near the start of Planchart's highly influential article *The Flower's Children*, he decided that, since he would be discussing clausulae in depth, he should define in a footnote what he understood by the term "clausula":

> A clarification of the terminology I will use in this study might be in order. Medieval authors such as Anonymous IV use "clausula" as a synonym of "punctum" — that is, a phrase […] — so a given clausula might be in discant, organum, or copula. Some modern scholarship not only restricts the term "clausula" to a phrase or section in discant, but even to those discant compositions copied separately from the body of the organum. For convenience's sake, I will restrict my use of the term "clausula" to indicate *discant* clausulae, whether they are embedded in an organum or copied separately, as most of them are in the clausula series of W_1 and F (Planchart 2003, p.305, emphasis original).

Planchart's other footnotes that define terms in this article, such as the footnote that directly follows the one quoted, discussing the 'ambiguous' term "organum" (Planchart 2003, p.305), are clear and to the point, but the above clausula definition is rather confusing. According to Planchart, medieval writers used the term "clausula" to refer to any section or phrase of music. However, modern writers use two separate definitions: the broader definition being a section in discant (generally understood today to be all voices moving in modal rhythm, with no sustained tones in the tenor),[1] and the narrower definition being that of a set of so-called "substitute clausulae" (free-standing separate sections of polyphony copied elsewhere in the manuscripts that were at times swapped in to replace their equivalent sections in the main body of organa).[2] Planchart elected here to use this term to refer only to sections of discant that swap or have been swapped, whether they appear in the substitute fascicles or exist already in the organa. He therefore purposefully left out two categories that would also have been considered to be clausulae, as per his interpretation of medieval usage: sections of organum purum that are replaced by discant "substitute" clausulae, and sections of polyphony — whether in purum or discant — that are never replaced. Planchart was therefore making

[1] See the definition given in Flotzinger, Sanders and Lefferts (2001).
[2] See the related footnote on p.10.

a bold proposition: that the general consensus on the definition of the term “clausula” is incorrect and incompatible with medieval usage. Despite this, he used this perceived faulty definition ‘for convenience’s sake’ simply because making a full argument for a broader definition of the term “clausula” would not only divert the topic of his study, but perhaps also confuse readers who were accustomed to the supposedly incorrect definition. However, such a claim, if true, is far too important to be left as a historical footnote. More than just simply a word, the definition of the term “clausula” circumscribes a broad swath of surviving thirteenth century music, and affects how all current study in this area is undertaken.

Bukofzer wrote in 1954 that ‘the history of the clausula is yet to be written’.[3] Although much research has been undertaken since then to account for the history of the clausula in the twelfth and thirteenth centuries, little work has been attempted to understand how terms such as “clausula” are used in modern academic literature. As the Planchart quote demonstrates, and as this chapter will argue, although the clausula is considered central to understanding musical reuse in the ND repertory, it has been used variably by scholars in the modern age as a historiographical touchstone for musical reuse, and altered to fit the prevailing arguments surrounding discant and substitution. This has had the effect of sidelining other forms of reuse, such as indirect concordances and melodic formulae.

As this study will be harnessing empirical and digital methodologies, it is vitally important that the terms used in the research are defined exactly. This chapter introduces the term “clausula” as the central study for how musical reuse has been, and continues to be, viewed in modern literature as the primary channel for musical reuse in the thirteenth century. It brings the terminological issues present in the clausula to the foreground and, by conducting a survey of literature concerning clausulae — such as research, music histories, and dictionaries of music — studies the clausula from a historiographical perspective through writers’ use of the term “clausula” as the locus of musical reuse. It will therefore bring together the narratives of how the term “clausula” was brought into modern study of the repertory, how it was introduced into music histories and dictionaries as the “missing link” between organa and motet, and how its definition was narrowed in scholarship into a particular musical form. In so doing, the chapter will highlight inconsistencies in these definitions, not only between medieval and modern usage, but also within the selfsame contemporary literature.

It is well known that the clausula constitutes only one (albeit the most visible) expression of musical reuse in the ND repertory (see previous discussion beginning p.6). This chapter argues that scholars’ past preoccupation with the “true” medieval

[3] ‘Die Geschichte der Clausula die noch geschrieben werden muß’ (Bukofzer 1954, p.563).

meaning of the term — such as used by Anonymous IV — has clouded the wider picture of musical reuse within the repertory.[4] Although use of the term necessarily began from a reading of medieval theory, what writers such as Anonymous IV really meant by the term "clausula" — if such a concept can be defined — is now less relevant to modern study than our relationship to how other modern writers define it, as this term is now used to reflect a modern rather than medieval conception of the ND repertory.

The focus here is therefore not on the medieval history of the clausula, but on its life in the nineteenth, twentieth, and twenty-first centuries. Nonetheless, the chapter will seek to create definitions of the term "clausula" that reflect the variety of modern usage. Scholars of medieval musicology, liturgy, and Latin lexicography have long been interested in the definition of the term "clausula" in the thirteenth century (and a discussion of this can be found in "2.7. The Anonymous IV passage"). However, the purpose of this chapter is not to uncover a "true" meaning, but to demonstrate the ways in which modern definitions have affected modern study. This is done so that this study can best relate the term "clausula" to the broader notion of musical reuse.

This treatment will seek to highlight how study of the clausula has, in the past, precluded the possibility of wider surveys of musical reuse, and that this is due not only to the clausula's immediately apparent manifestation within the sources, but also its vital historiographical importance for the development of core tenets of Western musicology, such as composition and authorship. In stark contrast to more nuanced forms of musical reuse such as indirect concordances and melodic formulae which work in opposition to such ideas, study of the clausula has pushed the conception of the ND repertory towards a closer — and perhaps more restrictive — link to static composition, that is, the versions of musical compositions recorded in writing.

It is an almost impossible task to extricate the idea of the clausula from that of the operational model of thirteenth-century music, as it was the substitute clausula that formed the base material for the large majority of early motets,[5] and the motet is a key form from which historical narratives of Western music are commonly strung. Therefore, in considering the context of the issues of clausulae outlined above and the possibility of the term "clausula" being defined variously across literature, this

---

[4] This chapter will make multiple references to what is perhaps the most famous passage of Anonymous IV. As Wegman (2015, p.695) observes, the passage has 'repeatedly been scrutinized for nuances and shades of meaning that might previously have been overlooked'. These repeated scrutinisations and disagreements in interpretation are part of the historiographical issue this chapter deals with in depth. Nonetheless, the most recent English translation of this passage is from Roesner (2001, p.227) and runs thus: 'And note that Master Leoninus was an excellent *organista*, so it is said, who made the great book of organum on the gradual and antiphonary to enrich the Divine Service. It was in use up to the time of Perotinus Magnus, who produced a redaction of it and made many better clausulae, that is, puncta, being an excellent *discantor*, and better than Leoninus was'.

[5] Even with the persuasive evidence of her own work on the subject of the complex and two-way relationship between clausulae and motet (Bradley 2013), Bradley is always careful to frame her examples as the exception, rather than the rule, and the process of clausulae becoming motet 'remains convincing in many cases' (Bradley et al 2019).

discussion is concerned less with clausulae per se as to their "true" definition, but rather with the concept of the clausula in the minds of modern writers: this is not the "clausula" but the "clausula-concept", the well known substitution narrative that links organa to motet (Flotzinger 1980).

The clausula-concept is seen as the bridge that links the "early" organum to the "later" motet, and it is tempting to see in the clausula-concept the idea that we have captured a snapshot of a musical revolution in progress. Substitute clausulae themselves provide evidence to support the theory that the Magnus liber evolved from a rhythmically-ambiguous set of extemporised two-part music into a repertory of tightly-controlled, rhythmically-defined, and premeditated composition through a mostly literate process of recomposition (Roesner 2001, p.265). The substitution process of clausulae, where both the earlier and later versions are transmitted in the same manuscript, allows us to feel as if we have caught Perotinus red-handed during his process of revision. At no other time in the next 500 years, i.e. not until the widespread preservation of composers' sketches and *Nachlass*, do we have such an insight into musicians' supposed process and, as a result, the clausula is fertile ground for analysis and discussion of music-making in the thirteenth century. It goes without saying, therefore, that the clausula-concept forms a central part of thirteenth-century musical exegesis. There are not one, but two slippery yet related historiographic concepts to catch hold of: the clausula defined as form and evidence of modernisation in the Magnus liber, and the clausula-concept of substitution.

2.1. Precursors to motet

For all early writers of music history in the eighteenth and nineteenth centuries, the genesis of the motet was of little importance. For Burney, the motet appeared fully formed in the fourteenth century in the works of Philippe de Vitry, but the motet was not worth discussing in any detail until the sixteenth century (Burney 1789). Fétis, writing almost a century later however, was one of the first to show some interest in the early motet and traced its beginnings to the 'end of the twelfth century'.[6] It was not until the end of the nineteenth century however, that writers had the necessary source materials to be confronted by the clausula's influence on the motet. However, the process was not yet understood. For example, Parry described the tenor of early motets as:

> A sort of nonsense part […] sung to nonsense syllables, such as "Balaam," or "Portare," or "Verbum," or "Angelus" […] the practice was so well understood that the composer merely wrote the word once at the beginning of the piece, and the

[6] 'Motets de la fin du douzième siècle' (Fétis 1869).

singers (generally those who took the lower part) fitted it in as seemed to them good (Parry 1893, pp.100–101).

In 1898, the link between organa and motet was made explicit by Meyer in *Der Ursprung des Motetts*. As part of explaining his discovery, Meyer created the very first iteration of the clausula-concept. He first identified that, within the polyphonic practice of the time, it was the melismatic sections of chant that were of greatest interest to thirteenth-century musicians, and Meyer anachronistically termed those sections "coloraturas" (*Coloraturen*). He described these coloraturas as the highlights and most virtuosic passages within the compositions, concluding that 'many composers only recomposed these 2–4 highlights of the individual antiphons, leaving the other parts aside'.[7] Meyer characterised the creation of clausulae fascicles as an organisational and notational shorthand rather than any musical choice designed to keep these sections separate: rewriting only that section of polyphony that was recomposed was done to save space and to prevent the collection from becoming untenable.[8] However, this process of choosing sections to sing 'according to need and personnel' was confusing to Meyer, but he did not pass judgement on it other than to term it 'a very unusual practice'.[9]

Meyer described the origin of the motet as a new invention of the twelfth century, prosulated motets, whereby clausulae were taken and new texts composed to fit the upper voice(s) (Meyer 1898, pp.310–311). The clausula-concept for Meyer was therefore a transitional form: sections of organa "coloraturas" were separately written, and it was these sections to which texts were added and clausulae transformed into motet. It can be assumed, therefore, that if the "coloraturas" Meyer defined are candidates for motet, then those sections of polyphony must have been those so-called "substitution" clausulae.

Such a clausula-concept was defined more succinctly by Ludwig in 1905 where, citing Meyer, he defined the genesis of motets as 'the passages where the liturgical melody is particularly lively, subordinated to the strict rhythm of the upper voice'.[10] Unlike Meyer, Ludwig appeared to be defining the clausula-concept as any section in discant. Therefore, for Ludwig the concept was not limited simply to independent

[7] 'Diese Coloraturen waren die Glanzpunkte, die Bravourstellen der verschiedenen Komposition; ja, mir ist wahrscheinlich, daß in der Zeit der eifrigsten Kunstübung viele Komposition überhaupt nur diese 2–4 Glanzpunkte der einzelnen Antiphonen neu komponirten, die übrigen Teile aber beiseite ließen' (Meyer 1898, pp.308–309).

[8] 'Wollte man jede komponirte Antiphone so oft ganz schreiben als verschiedene Kompositionen vorhanden waren, so ging der Umfang der Sammlung ins Ungehenerliche' (Meyer 1898, p.308).

[9] 'Konnte nun nach Bedürfnis und Personalverhältnissen diese oder jene wählen […] Die Komponisten besaßen eine ganz ungewöhnliche Übung' (Meyer 1898, p.310).

[10] 'Es sind die Stellen, an denen die liturgische Melodie besonders lebhaft melismiert verläuft, sich dem strengen Rhythmus der Oberstimme unterordnet' (Ludwig 1905, p.7).

sections transmitted in separate fascicles: although such passages were 'often composed in large numbers', Ludwig's wording does not exclude the possibility of such passages existing embedded in organa.[11]

Meyer and Ludwig's findings were then summarised in Leichtentritt's *Geschichte der Motette* (1908), where we can see the alluring narrative of the clausula-concept and its function as the vehicle for early motet beginning to form. Leichtentritt was creating a history of motet in all its forms, and constructed a narrative that depends on these passages being 'precursors' (*Vorläufer*) of motet (Leichtentritt 1908, p.4). Although neither Meyer nor Ludwig used this framing in their focused research on early motet, the implication created by Leichtentritt was that of a hierarchy, where the clausula-concept inhabits an earlier, inferior, or incomplete version of the motet. The value of the clausula-concept itself is minimised in favour of the "progressive" motet.

2.2. The term "clausula"

In 1864, Coussemaker published *Scriptorum de musica medivi aevi*, which included an edition of Anonymous IV, so called because it was the fourth anonymous text in Coussemaker's collection. Coussemaker remarked that this work was 'among the most outstanding works of measurable music of this age', but it had 'so far not affected the research of learned men'.[12] Meyer had used the words of Anonymous IV in his article, but had focused on the work of Perotinus in 'abridging and correcting' Leoninus' work rather than the method by which Perotinus was purported to have completed it.[13] Ludwig, however, began to lean upon the terminology "clausulae sive puncta", using Coussemaker's edition to explain the clausula replacement process outlined by Meyer. Ludwig defined the clausula-concept in terms of tenor rhythm, and a custom of composing discant not only once:

> Sondern die gleiche Tenormelodie rhythmisch verschiedenartig zu behandeln oder einer dieser rhythmischen Gruppierungen verschieden behandelte Oberstimmen gegenüberzustellen, also mehrere Kompositionen dieses Tenortextes zur Auswahl zu bieten, die abwechselnd an dieser Stelle zum Vortrag kamen, während die übrige Komposition des liturgischen Stückes zunachst blieb (Ludwig 1906, p.516).
> *But to treat the same tenor melody in different rhythms or to contrast one of these rhythmic groups with differently-treated upper voices, i.e. to offer several composi-*

[11] 'Es ist bezeichnend, daß solche Stellen in großer Menge auch ohne Widerholung des Übrigen vielfach komponiert und in besonderen Faszikeln gesammelt aufgezeichnet sind, ein Beweis für die starke Wirkung, die solche formal mehr geschlossenen Abschnitte innherhalb des breit ausgesponnenen Rahmens des Ganzen immer aufs neue hervorbrachten' (Ludwig 1905, p.7).

[12] 'Inter prestantissima hujus etatis de music mensurabili opera hoc quidem certe precipuum claret ac summum. Mirum adeo est quod, elsi apud Hawkins jam prius memoratum, eruditorum virorum indagationem hucusque haud commoverit' (Coussemaker 1864–1876, p.xx).

[13] 'Aber das Hauptverdienst hatte der berühmte Meister Perrotinus (magnus); er hatte des Leoninus liber organi gekürzt und verbessert' (Meyer 1898, p.323).

> *tions of this tenor text to choose from, which were performed alternately at this point, while the rest of the composition of the liturgical piece initially remained unchanged.*

Ludwig stated a similar view in 1909, that of the clausula-concept being the notational result of the act of Perotinus shortening the Magnus liber, and the earliest motets being ‘nothing more’ than clausulae with texts.[14] Just like Meyer, Ludwig’s emphasis was always on motet, and he drew a clear link from organa to motet using the clausula-concept. This is apparent in how Ludwig read Anonymous IV with a focus on ‘progressive development’ which led to an emphasis on the idea of Perotinus making “many better clausulae”, i.e. those that would lead to motet.[15]

Ludwig summed up his theories in what is generally thought to be the beginning of ND studies proper, his *Repertorium Organorum Recentioris et Motetorum Vetustissimi Stili* (1910). This was not only a complete catalogue of ND sources as they were understood at the time, but also a broadening and fleshing out of Ludwig’s clausula-concept to consider and encompass the entire repertory of organa. It is commonly thought that *Repertorium* (his culminating work) was the beginning of the clausula-concept,[16] but it was created in a pre-existing research context of discovering the origins of the motet. *Repertorium* is the filling-out of theory regarding early motet but, in completing a full study of early motets, Ludwig’s scope was necessarily broadened to also consider their sources in organa and substitute clausulae.

Once again, Ludwig acknowledged his borrowing of Anonymous IV’s terminology at the outset and used it throughout *Repertorium* to refer to the act of substitution (Ludwig 1910, p.15). Although Ludwig was not the first to use the term, he was the first to completely decontextualise the term “clausula” from the medieval context of Anonymous IV and reuse it in a new way. However, Ludwig’s definition is not as strict as we might like to believe. Rather than committing to the term “clausulae”, Ludwig also used the more general term *Teile* (passages) or *Ersatzteile* (substitute passages), as well as *Quellen* (sources), and this terminology aliasing caused his clausula-concept to become looser than its initial definition.[17] Indeed, Ludwig also used this very common term *Teile* to describe countless other ideas, such as parts of manuscripts or parts of tenor melodies. It is therefore difficult at times to understand exactly what kind of

[14] ‘Jedenfalls ist die musikalische Substanz der ältesten Motetten, die weiter nichts sind als Perotins „clausule [sic] sive puncta,“ mit neu gedichtetem syllabischem Text in den zur liturgischen Tenormelodie neu hinzukomponierten Oberstimmen’ (Ludwig 1909, p.207).
[15] ‘Es ist begreiflich, daß sich gerade in diesen Ersatzstücken die fortschreitende Entwicklung der Kunst des mehrstimmigen Satzes besonders lebhaft widerspiegelt’ (Ludwig 1906, p.516).
[16] For example, Mathias (2020, p.6) asserted that Ludwig, in *Repertorium* was ‘the earliest commentator to propose such a function’.
[17] Bradley (2018b, p.5) has alluded to the problematic way in which the word *Quellen* implies that clausulae were nothing more than material for creating motets.

"part" or "passage" Ludwig was referring to.

Although Ludwig defined the clausula-concept as above, using the narrative of organa to clausula to motet, he made explicit that his catalogue of clausulae contains only discant sections that have corresponding passages in the substitute fascicles, and whose tenors are also used as the basis for motets.[18] In other words, the clausula in Ludwig's conception is not simply the discant section that replaces another passage, but is also the passage that has been replaced. It is now well known that discant is more likely to replace another passage of discant than purum, but nonetheless there are a significant number of purum passages that are thus equated with the term "clausula" in *Repertorium*. For Ludwig, 'a small selection [of these] sufficed' (*genügte eine kleine Auswahl*) of these, and they are generally marked with "si|" for arrhythmic and "oγ" for irregular mode V tenors (Ludwig 1910, pp.15–16).

Despite this, and although equated with the clausula-concept, substituted passages are never called by the name "clausula": when Ludwig came to list clausulae, he listed only so-called "substitute" clausulae, and kept them separate from what they substitute, or have been substituted by. This perceived discrepancy was first highlighted by Dittmer nearly fifty years later in a footnote: 'one gains the impression, that Ludwig had recognised as a clausula only those passages of organum, which had counterparts among the series of substitute clausulae' (Dittmer 1959b, p.102). Ludwig's inclusion of purum in the clausula-concept has since been taken as an eccentricity of his work over more modern efforts (Smith 1964, pp.7–26; . What is more likely is that the catalogue's inclusion of purum — by the famously fastidious Ludwig[19] — is not a glaring error but in fact demonstrates his opinion that the clausula-concept is not the items themselves, but the process of substitution.

2.3. The clausula in music history

Ludwig's accomplishment of *Repertorium* was extremely influential, and his work has not yet been surpassed in depth or scope. Nevertheless, his primary concern was still motet: his study of organa and clausulae was undertaken to support his linear narrative of organa becoming motet. Although modern scholarship now views his work with full knowledge of his motet-centric bias, we must not forget the impact that Ludwig's work had on research into thirteenth-century music in the early twentieth century. For scholars after Ludwig, the most crucial items to discuss were those discant clausulae that became motet, as it was through those items that the "progression" narrative could be created.

[18] 'Diejenigen Partien, 1) in denen im T. [Tenor] ein bestimmiger sog. [sogenannter] „Modus" herrscht, 2) für die in den Ausschnitt-Faszikeln Ersatz-Kompositionen vorliegen und 3) deren T.-Melodie als Motetten-T. Verwendung fand' (Ludwig 1910, p.15).

[19] For examples of Ludwig's attention to detail, see Busse Berger (2005, pp.9–31).

The tidy narrative of organa becoming motet is quite an alluring story, and appears to be backed up by the Anonymous IV comment on “many better clausulae” (i.e. discant). As a result, the tight rhythmical construction of a “typical”, “Perotinian” clausula, with all voices moving in fast modal rhythm, caused the clausula-concept to not only form part of a repertorial solution but also become a rhythmic and stylistic discourse. Early musicologists such as Ludwig separated the stylistic concerns of purum versus discant from the act of swapping sections using the clausula-concept. However, three independent threads caused the issue of modal rhythm in discant to become partially conflated with the term “clausula”:

a) the discussion of the clausula-concept alongside discant;
b) a narrative impetus towards motet; and
c) the idea of Perotinus being both the agent behind clausulae and an “excellent discantor”.

For example, Apel, in his notational survey (1941), described four kinds of thirteenth-century notation to consider: ‘syllabic notation (simple conductus); duplum notation (organa dupla of the earlier, Leoninus period); modal notation (organa and clausulae of the Perotinus period); and motet notation (earliest motets)’ (Apel 1941, p.219). Note, only Perotinus is credited with modal rhythm and clausulae. The clausula-concept here becomes more than a method of swapping polyphonic sections around, and becomes the very genre that defines Perotinus’ output. This allows Apel to describe a *Benedicamus domino* as consisting of ‘a first section with widely spaced tenor notes, and of a second section showing continuous motion in the tenor. Sections of the latter type are called clausulae’ (Apel 1941, p.217). In other words, a “clausula” (and by extension the clausula-concept) in Apel’s definition is not only a structural choice but also a musical style, and therefore begins to encroach upon what would also be termed in Apel’s study “discant”. Discant and clausulae become a single concept: clausulae are sections in discant, but “substitute clausulae” are those sections contained in the independent clausulae fascicles (Apel 1941, p.215). The clausula-concept is narrowed to include only sections in discant, but simultaneously expanded to include any section that follows discant rhythm patterns, regardless of whether that passage is swapped among substitute clausulae. Ludwig’s original meaning of the clausula-concept as a process becomes limited to the “substitute” of “substitute clausulae”, and “clausula” alone is expanded to mean “discant”, such that the clausula-concept might as well be described at this point as “substitute discant”.

The mid-twentieth century was a pivotal time for the modern history of the clausula. The flurry of dictionaries and histories of music written in the postwar period began to pick up on the clausula-concept as the narrative tool that Ludwig had effectively designed it for. However, the intensely narrative focus of these new pedagogical tools helped to further blur the distinction between clausulae and discant. For example, Grout (1960, p.81) defined clausula and discant in no uncertain terms as synonyms:

> This new contrapuntal style in which all the parts are in measured rhythm came to be called discant; it did not exclude occasional short melismas, particularly at cadence points, but for the most part the two voices moved strictly according to the rhythmic modes. A section written in discant was called a clausula.

He framed the clausula-concept as a linear process, whereby ‘older clausulae were replaced with faster movements in definite and stylized patterns’ (Grout 1960, p.82), and elevated the substitute clausula to a genre in its own right as the precursor to motet: clausulae became ‘quasi-independent pieces’ that ‘eventually evolved into a new form, the motet’ (Grout 1960, p.81).

In other music histories, the term “clausula” subsumed discant completely. Hughes’ chapter on “Music in Fixed Rhythm” in the *New Oxford History of Music* (Hughes 1954, p.311–352) avoided the term “discant” entirely, preferring to use the term “clausula” exclusively. Although the original *Oxford History of Music* had quoted Anonymous IV at length and mentioned Leoninus and Perotinus, it did not link to motet using the clausula-concept, as the motet was introduced as an independent genre to be studied separately (Wooldridge 1901, p.339). However with Hughes the narrative force was very apparent, as he placed his subsection on clausulae at the very end of his chapter, the following chapter being concerned with motet. Indeed, Hughes’ focus on reaching substitute clausulae (and therefore motet) was made explicit: his introduction to organum began by announcing that its importance is ‘not for what it is as for what it produces later by way of the clausula’ (Hughes 1954, p.343). A byproduct of the term becoming so variable is that its definition also became dynamic, able to fill the hole left by other terminology. Hughes’ reluctance to use the term “discant” caused the term “clausula” to become more far-reaching than the definitions used by other writers.

An exception to this can be seen in Bukofzer’s article “Discantus” in *Musik in Geschichte und Gegenwart* (MGG) (1954) which has arguably stood the test of time, despite being nearly 70 years old. Smith (1964), writing just after the mid-century proliferation of dictionaries and histories that began to include the clausula-concept, described it as ‘one of the best and most comprehensive treatments of the clausula’ (Smith 1964, p.3). More recently, Mathias (2020) used the article as the starting point for his review of the clausula literature, and described how the need for further research identified in Bukofzer’s article had not yet been satisfied (Mathias 2020, p.16).

This is perhaps due to the fact that Bukofzer’s MGG article leaves many of its questions unanswered and does not preclude any possibility from the repertory. For example, the article begins by admitting that, for discant, 'the terminology of the medieval theorists is neither clear nor consistent',[20] and the entry maintains this equivocal stance throughout, rarely defining anything so narrowly as to be an easy target for objection. Bukofzer used discant to define clausulae but did not conflate them, defining discant first as the opposite of purum, but not going so far as to say that all discant sections are examples of clausulae (Bukofzer 1954, p.560). By approaching the subject this way, he managed to separate the historical concerns of the clausula from the musical concerns of discant. For Bukofzer, discant was reserved for most (but not all) sections with melismatic tenors, purum being used mainly for syllabic tenors. The clausula is a separate concept to discant, but need not be separated in the manuscript (i.e. into so-called "substitute clausulae"): they are simply certain sections in discant with melismatic tenors.

Bukofzer again defined the clausula in a dedicated section as 'that part of an organum of the Notre Dame school that is set rhythmically rather than in an organal sustained tone style'.[21] Once again though, Bukofzer carefully avoided overstating the relationship between discant and clausulae. He had defined discant as 'note against note' (*Note gegen Note*), but had also realised that by this definition most clausulae would not fit, as they 'only rarely go strictly note against note (homorhythmically) over long stretches', as in the manner of conductus.[22] Despite this, Bukofzer still categorised clausulae as a general form of discant. What we can infer from this is that, for Bukofzer, the clausula was something approaching a subcategory of discant: in modal rhythm but not exactly note against note.

On the subject of the clausula-concept, Bukofzer offered only a précis of Ludwig's views, but made an important point relating to chronology and Ludwig's biases:

> Infolge der Quellenlage ist es unmöglich, einen klaren Strich zu ziehen zwischen einer originalen Clausula und einem Ersatzteil, da die Clausulafaszikel nicht nur die ausgeschiedenen altmodischen Teile, sondern auch die im modernen Stil nebenund durcheinander überliefern. Eine originale Clausula wird ja dadurch, daß sie dem Ersatzteilfaszikel überwiesen wird, selbst zu einem fakultativen Ersatzteil, und somit verliert die Unterscheidung praktisch ihren Sinn […] Ludwig, der sein Augenmerk besonders auf die Motette gerichtet hatte, hat die Clausula weniger als selbständige Form denn als Vorstufe zur Motette angesehen (Bukofzer 1954, pp.562–563).
>
> *Due to the source situation it is impossible to draw a clear line between an original*

[20] 'Die Terminologie der ma. [mitteraltlichen] Theoretiker ist allerdings weder klar noch konsequent und kann nicht historisch auf einen Nenner gebracht werden' (Bukofzer 1954, p.560).
[21] 'Unter Clausula versteht man denjenigen Teil eines Organums der Notre Dame-Schule, der nicht im organalen Haltentonstil, sondern rhythmisch gesetzt ist' (Bukofzer 1954, p.562).
[22] 'Wenn auch die St. [Stimmen] der rhythmisch streng gemessenen Diskant-abschnitte nur selten über lägere Strecken streng Note gegen Note (homorhythmisch) einhergehen, so entspricht doch zumindest einer Note des T. [Tenor] eine Notengruppe in der Oberst. [Oberstimme]' (Bukofzer 1954, p.562).

> *clausula and a substitute passage, since the clausula fascicles transmit not only the discarded old-fashioned passages, but also those in the modern style side by side and mixed up. In fact, by being assigned to the fascicle of substitute passages, an original clausula becomes itself an optional substitute, and the distinction thus loses its practical meaning […] Ludwig, who had focused his attention particularly on the motet, regarded the clausula less as an independent form than as a precursor to the motet.*

To my knowledge, this is the first time that Ludwig's biases towards motet were questioned. Bukofzer correctly saw that the clausula as an independent form had been neglected in favour of its role in the clausula-concept, and the siren song of the "progressive" motet had caused the clausula itself to be overlooked. Simultaneously, Ludwig had glossed over the difficult problem of chronology and not considered that a more "modern-looking" clausula may in fact be much older.[23] What we can conclude is that somewhere between Ludwig's *Repertorium* and the mid-century histories and dictionaries that digested it into the clausula-concept narrative, Ludwig's original clausula-concept of the process of making motet was merged with its most common manifestation in the discant substitute clausula.

2.4. Study of the clausula

Therefore, one of the most interesting trends emerging from mid-century articles such as Bukofzer's was the tendency to take the clausula-concept as read and focus less upon the act of substitution and more upon the discant clausulae (particularly substitute clausulae) themselves as independent items that can be considered apart from the source organa. This is the elevation of the clausula from simply "substitute passages" into an independent genre.

As mentioned previously, a core component of the clausula-concept is the narrative of organa transformed to motet through clausulae. This primacy of the motet can be seen in Ludwig's study and then carried over almost completely unchallenged into dictionaries and music histories. This narrative frames the clausula as fundamentally a transitional form: although it had been elevated into consideration as an independent genre by writers such as Apel, Hughes, and Bukofzer, its characterisation as a style rather than a section was created as a means to the understanding of early motet. Within the research context of the early twentieth century, focused almost entirely on motets, the study of clausulae outside of the clausula-concept was unfeasible. Clausulae —

[23] Handschin had already obliquely warned against this idea of medieval progressivism: 'A monument that is "advanced" in one respect may be "backward" in the other' (*Ein Denkmal, das "fortgeschritten" in der einen Hinsicht ist, kann in der anderen "zurückgeblieben" sein*) (Handschin 1930, p.52).

however they were defined — were studied in the early twentieth century simply as the tool by which a medieval musician might create a motet, without much thought for the purpose of clausulae themselves.

2.4.1. Clausula without motet

Isolated study of clausulae arguably began with William Waite's doctoral thesis, which was later published as *The Rhythm of Twelfth-Century Polyphony* (1954). This may seem like an odd place to start, as Waite's study does not deal with clausulae directly or in any detail, either as sections or substitute discant, and Waite's argument that even purum was organised in strict, rational rhythm is now extremely unfashionable in many academic circles. However, this was only part of his conception of the music of the ND repertory, and Waite's study is perhaps the first *prima facie* study that focused on organum itself and not motet. It was therefore through this route that the study of clausulae outside of the confines of the organa-to-motet narrative truly began, around the same time that interpretations of the original motet-centric clausula-concept of Ludwig were being disseminated through music dictionaries and histories.

However, with the benefit of 50 years of study on the topic already available to him, Waite could take the clausula-concept as read, and focus on the "how" and "why" rather than the "what" of Ludwig's study. Crucially, Waite avoided motet and *cum littera* notation, and focused entirely on *sine littera* and modal rhythm. The study is therefore unhindered by any previous focus on clausulae as proto-motet: Waite dealt with the clausula-concept in its original, thirteenth-century context of expanding the repertory of the Magnus liber.

Depending on the needs of this context, Waite's definition of the term "clausula" fulfils multiple ideas and retains its dynamic status as a catch-all term for the gaps between organa and motet. He began by translating the famous Anonymous IV passage using the phrase 'substitute sections' to explain the term "clausula" and the work of Perotinus (Waite 1954, p.3). Later, he invoked the clausula-concept to describe the 'large group of smaller compositions, new versions of sections of the organa designed to replace equivalent sections of Leonin's Magnus liber' as 'the "clausulae sive puncta" which Anonymous IV attributes to Perotin' (Waite 1954, p.6). Two pages later, he defined clausulae as 'sections of the organum where the tenor itself has a melisma', and where 'Leonin speeds up the rhythmic progression of the tenor' (Waite 1954, p.8). Variously then, Waite's usage of the term fulfilled that same one-size-fits-all dynamic purpose, being simply a section or, more particularly, a passage with a modal tenor, especially those that replace older polyphony.

From this study, Waite's article *The Abbreviation of the Magnus Liber* (1961) emerged, and this is to my knowledge the earliest attempt to trace the history of the clausula, both in the medieval and in the modern world. Waite asked directly 'what does Anonymous IV mean by "clausulae or puncta?"' and went on to attempt to answer his own question (Waite 1961, p.149). He did this by tracing the etymology of the term from Latin grammarians, and its movement from denoting the end of a phrase to denoting the phrase itself. In so doing, he accepted the premise that a clausula originally meant in grammar simply a phrase. However, according to Waite, the term 'was taken over by the early theorists of rhythmic poetry' in the twelfth century (Waite 1961, p.149), and so by the time Anonymous IV was writing in the late thirteenth century, he 'must have had the theory of rhythmic poetry in mind' (Waite 1961, p.150).

2.4.2. Clausula style

In complement to Waite's theories was Dittmer's *A Central Source of Notre-Dame Polyphony* (Dittmer 1959b). Although ostensibly concerned with the "reconstruction" of part of an ND source from the various fragments from Johannes Wolf's private library (**MüA**), Dittmer expanded his remarks in numerous long footnotes concerning related matters, including clausulae. As previously mentioned, the most interesting passage for this particular discussion is Dittmer's critique of Ludwig's clausula selection (Dittmer 1959b, p.102). Dittmer was at that time also engaged in preparing an edition of Ludwig's *Repertorium* for publication, and we can see here Dittmer struggling to reconcile Ludwig's process-based clausula-concept with the sectional conception in place in the mid twentieth century. Dittmer realised that simply compiling a list of clausulae from Ludwig's work was impossible, and forced a redefinition of the term:

> In this study, we conceive of a clausula as something musico-technical, rather than historical: an organal section is a part of an organum, in which the upper voice presents an extended melisma over tones held in the tenor, whereas a clausula is a composition either inside or outside of an organa, in which the tenor is arranged rhythmically (Dittmer 1959b, p.102).

This goes straight to the heart of the terminological difference. For Ludwig, the clausula-concept was a progression of the replacement of "old" organa by "new" discant, and a "list of clausulae" — if such a list could be created comprehensively from Ludwig's work — would reflect that narrative. Dittmer, however, defined the term "clausula" in a way much closer to how it began to be seen in mid-twentieth-century scholarship: as a section that uses a particular (discant-like) polyphonic style. Dittmer then catalogued the differences between a hypothetical Ludwig catalogue and the catalogue of clausulae by his own reckoning, thereby providing the clearest definition thus far of exactly what a writer should mean by the term "clausula": one member of an exhaustive catalogue. Dittmer's catalogue (1959b, p.105, p.124) reduced the question to a simple statement of set membership.

2.4.3. Clausula as section

This re-evaluation of the clausula-concept fed directly into Smith's study of the clausula (Smith 1964). This was supervised by Waite (University of Pennyslvania 2014), and Smith directly cited Dittmer's study and list of clausulae as the starting point for his catalogue (Smith 1964, p.13). He rightly saw that 'various approaches to a study of the clausula led to an early, and increasingly strong, realization that the clausula lacked a firm foundation in repertorial studies such as that which its offspring, the motet, was provided by Ludwig's *Repertorium*' (Smith 1964, p.4), and the present chapter has described the variety of approaches that led up to this moment. Instead, Smith aimed to create a repertorial study of the clausula at the same level that *Repertorium* had provided for motet, and he considered his thesis 'groundwork' for piecing together the story of clausulae later (Smith 1964, p.5).

Smith echoed Dittmer's concerns over Ludwig's catalogue, but went into more detail for his reasoning, accepting the previous focus on motet as a possible reason for a lack of clausula study (Smith 1964, p.7). Smith was not as generous to Ludwig as Dittmer however, considering Ludwig's principles in clausula selection as 'unsystematic and very incomplete' (Smith 1964, p.10), concluding that 'Ludwig's selection and numbering of discant clausulae of organa are based upon principles which are not compatible with a systematic and exhaustive cataloguing of *all* discant clausulae' (Smith 1964, p.12, emphasis original). As a result, Smith discarded Ludwig's catalogue and began again using a more sectional and tabular methodology.

This emphasis on sectionalisation, and the division of settings into their constituent passages, forms part of a trend that can be seen in this period of ND studies and still can be detected in some modern scholarship.[24] The clausula-concept had moved away from an idea of replacing some sections with new sections in order to improve, enlarge, or abbreviate the Magnus liber and was now pointed towards a formalisation of ND polyphony as consisting of a number of different sections that could be divided for analysis. This is not a purely objective process, and Smith admitted that 'there is more than one way of interpreting such matters' (Smith 1964, p.13). Smith's main discussion is fairly short, but it is in the second volume of the thesis that the sectionalisation of organa becomes apparent, in the diagrammatic analysis that divides each organa into its component sections, depending on which source contains that section and whether that section is in purum or discant.

[24] See, for example, the definition given in Burkholder et al (2006) on p.61.

Extract from Smith's diagram of O1

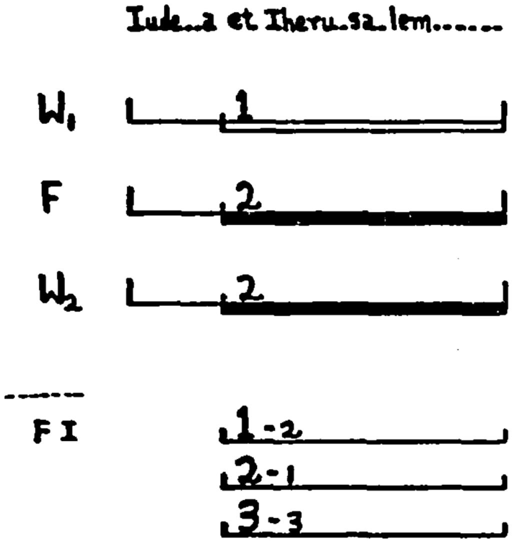

Fig. 2.1

2.5. The "substitute" clausula

Despite the criticisms Smith levelled at Ludwig, he still organised his diagrams by their Ludwig numbers. Each diagram presents the organum as a series of sections, alternating purum and discant, and the various sources for each setting are stacked one on top of each other, such that a reader can see at a glance which sections take part in the clausula-concept, whether those clausulae are embedded in other organa, or whether they exist in substitute fascicles. To take the first example, *Iudea et Iherusalem* (O1), from Smith's diagram it can quickly be seen that the word *Iudea* is set identically in $\mathbf{W_1}$, **F** and $\mathbf{W_2}$, but for *Et Iherusalem*, $\mathbf{W_1}$ has one discant section whereas **F** and $\mathbf{W_2}$ have another (see Fig. 2.1). Furthermore, **F** has three "substitute clausulae" for this section. The diagrams show not only changes in discant, but also diverging sections of purum. For example, *Tamquam sponsus* (O2) has famously many different "substitute clausulae" for the word *Tamquam*, but what is less well known is that the following purum *Sponsus dominus procedens* is radically different between the $\mathbf{W_1}$ and **F** settings (see Fig. 2.2). However, such differences can instantly be recognised using Smith's graphical methodology.

Smith's diagrams contain a wealth of information for investigating the construction of organa. Various hypotheses about organa can be quickly backed up with evidence, or counterexamples found. For example, when considering the chronology that Ludwig set forth, that of $\mathbf{W_1}$ being the earlier, "Leoninian" Magnus liber and **F** and $\mathbf{W_2}$ containing the later, "Perotinian" version,[25] a quick scan of Smith's diagrams yields counterexamples such as O29 (*Regnum mundi*), where discant sections in $\mathbf{W_1}$ are

[25] 'Ludwig saw the states in the three major collections as standing in the chronological order W_1–F–W_2' (Roesner 1981, p.368).

Extract from Smith's diagram of O2

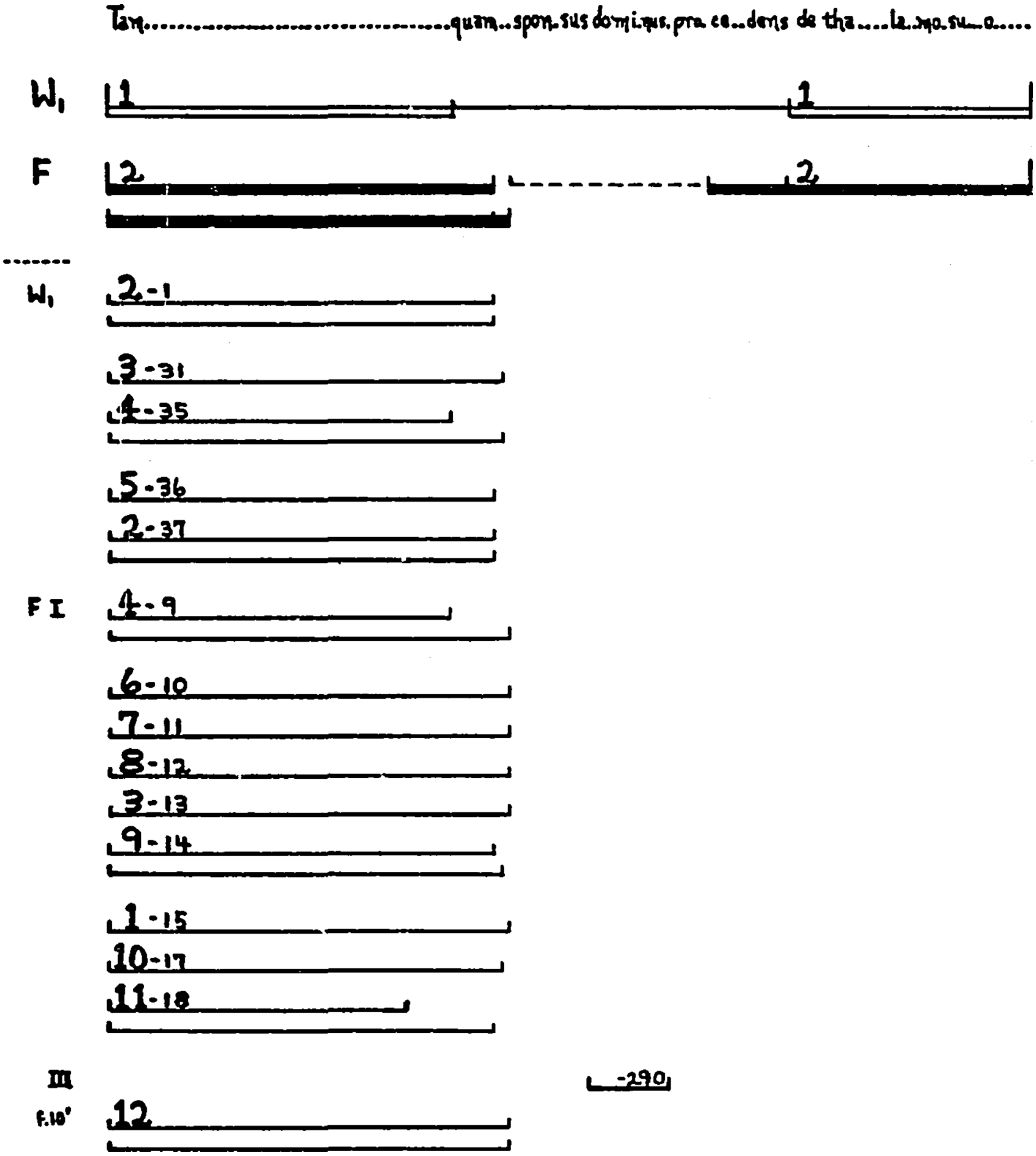

Fig. 2.2

"replaced" by purum in both **F** and $\mathbf{W}_2$ that did not exist in $\mathbf{W}_1$. If the "Perotinian" Magnus liber was a rhythmic and rationalising impulse, then such examples should not exist: in progressive terms, the music has "regressed" in style. In Smith's study, we see the confluence of two broad trends that then become the status quo: the narrowing of the clausula-concept to the discussion of "clausulae" rather than the process of substitution, and the sectionalisation of organa into component passages.

Since Smith (1964), a variety of full length studies have continued this trend of independent and sectionalised clausula study. Among these, we can count Flotzinger

(1969), who divided the clausula repertory by transmission status; Baltzer (1974), who further divided the repertory of clausulae into groups based on their tenor rhythms; and Mathias (2020), who focused on how such clausulae were created. For our present historiographical inquiry, however, these studies bear little relevance here as their research follows the same line as Smith's, that is, of the clausula as a genre and repertory somewhat independent from organum and not entirely reliant on motet. These studies conceive of the clausula repertory not as a precursor to motet, as Ludwig did, but as a partially self-supporting genre with intrinsic compositional processes distinct from organum as a whole. As a result, musical reuse — the primary focus of the present study — is an ancillary concern to these studies of the clausula repertory.

By the middle of the twentieth century, the definitions of the term "clausula" that focused on the clausula-concept — that is, the replacement of earlier, "Leoninian" purum and discant by more modern "Perotinian" discant — had given way to a dynamic definition of the clausula that prioritised its status as a section in strict modal rhythm over other considerations, at times equating this rhythm with discant, and at others creating the new independent genre of the clausula in the substitute fascicles. However, this later twentieth century study of clausulae as independent sections apart from the broader concerns of organum led some research further away from musical reuse through the clausula-concept and towards the study of clausulae themselves as independent and detached items.

Much of this movement can be attributed to Sanders, who believed that the substitution clausula-concept was not as convincing as other writers made it out to be, positing instead that the genre of the independent clausula was more likely cultivated by its own culture of composition. Although Sanders had earlier defined the clausula as 'the melismatic portions of responsorial chants and the *Benedicamus Domino*' (Sanders 1964, p.267), as early as 1967 he wrote that the clausulae were 'the first repertory of self-contained and independently shaped pieces of music with a cantus firmus' (Sanders 1967, p.243), and therefore clausulae were their own pieces that enjoyed a life outside of the clausula-concept. As evidence for this, Sanders noted how rarely clausulae actually found their way into "later" versions of the Magnus liber: only a third of "substitute clausulae" in $\mathbf{W_1}$ found their way into the **F** Magnus liber, and only 5% of **F**'s collection into $\mathbf{W_2}$'s Magnus liber (Sanders 1973, p.506).

In terms of definition, Sanders hinted in a footnote to the confused definitions of the term "clausula", but dismissed the idea that purum could be termed a clausula due to the proximity of the term "clausula" to discant in Anonymous IV:

> Given the evolutionary circumstances, present-day terminology ought not to be any less precise than it seems to have been in Perotinus' day. Since Anonymous IV

> reports that Perotinus «fecit clausulas sive puncta plurima meliora quoniam optimus discantor erat», it is evident that organal passages are not clausulae (Sanders 1973, pp.503–504).

Sanders equated the term “clausula” with only the substitute clausulae: $\mathbf{W}_1$ substitute clausulae were inserted into the organa of **F** and $\mathbf{W}_2$, and substitute clausulae from all three sources were used as sources for motet (Sanders 1973, p.505). In minimising the effect of the clausula-concept, Sanders created a new taxonomy whereby there are two styles: organal style (purum sections) and discant style (discant sections). At the same time there is an independent genre of discant that emerged out of organa, with links to motet, called “clausulae”. These are the compositions preserved in separate fascicles of the source manuscripts.

Sanders was responsible for the “Medieval” section of the updated article on “motet” in the *New Grove Dictionary of Music and Musicians* (1980), and worked his exclusionary definition of “clausula” into this article. Sanders defined the clausula here as ‘detached discant sections’ (Sanders 1980, p.617), and this definition gives the impression that clausulae are identical to discant sections in all but their placement in the manuscripts. He argued that, unlike discant sections embedded in organa, separate clausulae are more akin to motets, and motets were only understandable if paired with their source clausula. Using this line of reasoning, Sanders asserted that ‘it seems difficult to imagine the composition or the rehearsal of a motet without the aid of a melismatic model’ (Sanders 1980, p.617).[26]

The *New Grove* article on “clausula” however, was written by Flotzinger (1980). It was already clear from Flotzinger’s own study of the repertory that his definition of the term “clausula” followed this same distinction between discant and substitute clausulae (Flotzinger 1969). Whereas Smith’s study of around the same time was a study of “clausulae” being any discant part swapped between the sources (Smith 1964), Flotzinger’s study purported to catalogue “discant clauses” (*Discantussatz*) instead, reserving the term “clausula” for the so-called “substitute clausulae”:

> Obwohl in der Literatur meist die beiden Begriffe *Discantuspartie* und *Clausula* als identisch und austauschbar verwendent werden […] sollen sie hier auf ihre jeweiligen Überlieferungsformen (einerseits eingebettet in der Magnus-liber-Fassungen und andrerseits als Einzelstücke in eigenen Faszikeln gesammelt) beschränkt werden (Flotzinger 1969, p.16).
>
> *Although in the literature the two terms “discant part” and “clausula” are mostly*

[26]This has been roundly rejected by Smith (1989, p.144) and Bradley (2018a, p.55) by the simple fact that the rhythms of many motets have been “worked out” without the aid of such models.

> *used as identical and interchangeable […] they should be limited here to their respective forms of transmission (on the one hand embedded in the Magnus liber versions and on the other collected as individual pieces in their own fascicles).*

Therefore, when Flotzinger came to write his *New Grove* article, he had the same definition in mind: that of independent discant parts copied in separate fascicles. Flotzinger defined "clausula" etymologically, leaning on a definition of the term "clausula" that emerges from its use in Latin grammar and rhetoric as 'either the concluding of a passage, or the passage itself concluded' (1980). He therefore gave a similarly ambivalent musical definition of six historical meanings (loosely limited to 'the 11th to 15th centuries') and one 'formal sense' for the ND clausula repertory. In so doing, Flotzinger appeared to accept the same premise that Planchart (2003, p.305) described at the beginning of this chapter: that the word used in academic literature 'is reserved nowadays almost exclusively for […] the clausula of the Notre Dame period' (in this definition, substitute clausulae) (Flotzinger 1980), but in the medieval period it may have had any number of meanings. This is not to say that this approach is not without its weaknesses, however. By avoiding answering the question "what is a clausula?" directly, Flotzinger never truly defined what is and what is not a clausula in the context of the ND repertory, and he necessarily focused instead on an extremely concise and conservative narrative of the clausula-concept. For example, the definition claims that the purpose of clausulae in the ND repertory was to replace purum by discant, despite it being well known by that time that this narrative, although compelling, is only partly true, and substitute clausulae more often than not replaced other *discant* passages rather than purum.[27]

Such definitions, emphasising the independent and "compositional" aspects of substitute clausulae over their function as substitutes for passages of organa, allowed these clausulae to be studied and edited outside of their organal contexts. Whereas Tischler's edition of the *dupla* (1988) attempted comparison between all transmissions, vertically stacking clausula transmissions and centonising organa so as to compare like-for-like as much as can be possible, the editors of the *Magnus Liber Organi* editions believed instead that a critical edition was 'impractical and in large measure undesirable' (Roesner 1993, pp.lxi), and cast doubt upon the value of a comparative edition such as Tischler's, at least for the *dupla*. As a result, a significant portion of this edition is taken up with the many substitute clausulae in **F** and $\mathbf{W}_1$: one entire volume (Baltzer 1995) as well as large parts of two others (Roesner 1993, beginning at p.30; . It is clear that such a presentation would not have been possible if, at that time, clausulae

[27] See, for example, Waite (1961, p.151).

were still viewed as precursors to motets or inseparable from organum. The decisions that led towards a diplomatic treatment of the substitute clausulae in edition are intrinsic to the research context that allowed such a view of the substitute clausula to be presented in editions. It seems unfeasible that this approach could have taken place a few decades prior, when discussions about how a critical edition of the ND repertory could be constructed were still seriously considered (Karp 1966).

This approach is not infallible, however, and the trade-offs were pointed out both by Rankin (1997, p.140) who described the edition's organisation as a 'bold choice' that would require 'an abundant supply of bookmarks', as well as Bell (2006, p.312) who critiqued the lack of consistent editorial oversight between volumes. Although such a presentation allows a reader to investigate historical layering as it is presented in the sources, this arrangement also minimises the impact of the clausula-concept by retaining the codicological, rather than the organal, context of the substitute clausulae. Baltzer, the editor of the volume dedicated to the clausulae transmitted in **F** (volume five), noted four speculative uses for the clausulae: the familiar substitution and motet source reasons are listed first, but she also named 'clerical chamber music' and 'compositional études' as possible uses (Baltzer 1995, p.xxxix). Presenting clausulae separately as independent works doubtless elevates the last two uses (for which there is little evidence) at the expense of the first two (without doubt their primary function),[28] further separating clausulae from their organal contexts.

2.6. Alternate definitions

As this chapter has shown thus far, the clausula-concept has been an integral component of the dominant narrative concerning musical reuse in thirteenth-century music, from the late nineteenth century to the present day. This emphasis on the clausula has largely been the product of a historiographical impetus by scholars of the late nineteenth and early twentieth century, who in part aimed to construct the basis for Western art music in ND polyphony by identifying the origin of the motet in prosulated substitute clausulae. However, it has also been the case that scholars have sought to study substitute clausulae as an independent repertory of cultivated, literate polyphony. Regardless of the extent to which we believe the veracity of this claim, such treatment has served to elevate the "compositional" aspects of ND clausulae at the expense of their contextualisation within organum as tools for musical reuse.

However, not all definitions given in academic literature neatly fit into this motet- and independence-focused characterisation. As Planchart's footnote exemplifies,

[28] Some of this limited evidence for the independence of substitute clausulae from an organal context is collated in Tischler (1969) and Bradley (2018b, pp.247–257), the most striking being the clausula on the retrograde *Dominus* (**F**, f.150v).

beneath this narrative of the clausula-concept there is also an alternate strand of definitions that have taken a critical view of the accepted wisdom that the clausula is an object distinguishable from organum by either musico-technical or codicological criteria. Rather than those definitions that have focused particularly on stylistic markers such as discant for defining clausulae — the technical aspects of which have necessarily shifted focus away from the processes of musical reuse that were likely not driven by technicality — these alternate definitions, although far fewer in number, have greater applicability to this chapter's focus on contextualising how we view musical reuse.

The earliest example of such an alternate definition is Ursprung (1931), who equated the term "clausula" with "compositional sections" and, although he concurrently discussed the difference between purum and discant, did not place the clausula-concept as synonymous with either style:

> Diese Technik und solchen Tonsätzen bzw. Kompositions-„Abschnitten" („clausulae" oder „puncta") verbleibt der alte Name „Organum" oder sie werden nunmehr auch „Organum purum" geheißen […] wo aber die gegebene liturgische Melodie selbst schon reicher melismatisch ist, werden im mehrstimmigen Satz vielfach beide Stimmen straffer rhythmisiert, d.h. es wird die „modale" Rhythmik in Tenor und Duplum eingeführt, die Stimmen werden in ihrem rhythmischen Verhältnis systematischer gestaltet, die Abschnitte oder Clausulae gelangen zu eindrucksvollerer musikalischer Geschlossenheit. Für diese Satzweise bzw. diese Partien gilt der Name „Discantus" (Ursprung 1931, p.122).
>
> *This technique and such musical movements or rather compositional "sections" ("clausulae" or "puncta") retain the old name "organum" or, as they are now also called "organum purum" […] but where the given liturgical melody itself is richer in melismatic form, both voices in the polyphonic setting often have a tighter rhythm, i.e. the "modal" rhythm is introduced in tenor and duplum, the voices are structured more systematically in their rhythmic relationship, the sections or clausulae arrive at a more impressive musical unity. The name "discantus" applies to this type of movement or these parts.*

The tension between this definition and the discant-based definitions can be seen in Reese (1940, p.298), where he first defined clausulae as 'the larger sections into which the alteration of these styles [purum and discant] divides a composition', and then used the term "clausula" to describe passages in either style. In a footnote however, Reese made plain the divergence of the terminology:

> The term clausula is used by several writers, notably Ludwig […] for the discant sections only. It is applied in Ursprung [quoted above], […] however, to sections of

> both types. The passage from Anonymus IV[29] […] does not exclude either interpretation (Reese 1940, p.298).

There is one small exception to this: that as already made clear, Ludwig's clausula-concept did not exclude purum, but rather minimised its effect in favour of the motet. Ludwig's priority in dealing with discant is clear, as when he wrote about and used the term "clausula", he nearly always meant substitute discant. However, collating Ludwig's evidence together is not as simple as collecting his disparate catalogue into a "list of clausulae", as those who have attempted to do so have found out. For example, van der Werf, attempting to define what Ludwig meant by "clausula", had to settle for the unsatisfactory 'pieces that look like excerpts from organa' (van der Werf 1989, p.8).

Another alternate definition is given by Yudkin (1989, p.382), who avoided the clausula-concept entirely and defined the clausula simply by its etymological roots:

> The word clausula, like many terms in medieval music theory, is taken from Latin grammar, where it means a "clause" or part of a sentence. In the same way, the musical clausula is a part or section of a complete composition.

There is no mention of style here, only sections. Using this alternate definition, Yudkin created a new narrative quite different from the received idea of clausulae being immediately useful for motet: 'at first clausulae had no independent existence. They were exactly what they were called: "clauses" — parts of a whole […] clausulae gradually evolved from fragmentary, incomplete sections of a piece to entirely independent compositions in their own right' (Yudkin 1989, p.391). This is a different story to the usual idea of clausulae being created that were then used as sources for motet, and reframes Sanders' dismissal of "early" clausulae as in fact a positive aspect of them being "parts of a whole" (Sanders 1973). Yudkin instead believed that "early" clausulae were likely not suitable for texting, and it was only "later" clausulae (i.e. in discant) that were considered as independent and texted into motets.

Some of the more recent updates to the history of music that began as Grout (1960) have rewritten the definition of clausula as not contingent on discant, as was true up to and including the sixth edition (Grout and Palisca 2001, p.80), but as sections, perhaps beginning in purum, that nevertheless were a force for the creation of more discant:

> Anonymous IV writes that Pérotin edited the Magnus liber and "made very many better clausulae". By *clausula* the Latin word for a clause or phrase in a sentence, he meant a self-contained section of an organum, setting a word or syllable from the

[29] In modern Anglophone scholarship the treatise has been respelled as "Anonymous IV", but Coussemaker originally titled it in Latin as "Anonymus IV". See Taruskin (2005, p.173).

> chant and closing with a cadence. Since Léonin's organa consisted of a series of such sections, it was possible for Pérotin and others to write new clausulae designed to replace the original setting of a particular segment of chant. Typically, these new clausulae, known today as *substitute clausulae*, are in discant style, reflecting a growing preference for discant (Burkholder et al 2006, p.100).

A final author to discuss in this context is Roesner. Despite his extensive writing on the subject, Roesner was careful never to define the term "clausula" exactly.[30] Indeed, Roesner's work appears to assume that the reader already understands what is meant by the term "clausula", and so it is unsurprising that Roesner often qualified the term "clausula" by its style: "discant clausula". What is surprising, however, is the informal manner with which Roesner introduced the term 'organum purum clausula' (a term of which I can find no earlier instance in academic literature) in a figure caption as if it were already commonly used (Roesner 1979, p.182). In an academic context that appeared to define the term "clausula" as either synonymous with discant or as substitutes, Roesner was apparently happy to discuss 'both sustained-tone [purum] and discant clausulae', as well as a 'clausula in organum purum' (Roesner 1981, p.366, p.385).

For example, in Roesner's most comprehensive treatment of the clausula-concept (2001), he used the term "clausula" with an apparent freedom not seen before. His first mention of clausulae is at the opening of his discussion of *Alleluya, Adorabo*: 'F includes several discant clausulae where W_1 has organum purum' (Roesner 2001, p.238). This is indeed true. **F** and $\mathbf{W_2}$ include independent clausulae for "rabo" and "sanctum tuum" where $\mathbf{W_1}$ is in purum until "et confitebor". He also mentioned the 'W_2 clausula on "-torium" in *Alleluya, Posui*' (Roesner 2001, p.238), which has discant clausulae in all three main sources. However, he then moved into a discussion of 'the clausulae on "nomini" and "electum"' in *Alleluya, Adorabo* and *Alleluya, Posui* (Roesner 2001, p.241). Roesner was here discussing final purum passages present in some sources. Neither "nomini" nor "electum" appear in any collections of independent clausulae and are not catalogued by any other writers as clausulae, "nomini" appearing in $\mathbf{W_2}$ only and "electum" only as part of the organa in $\mathbf{W_1}$, **F**, and $\mathbf{W_2}$. The definition of "clausula" that Roesner was using here must have been slightly different.

In the same article, Roesner then analysed a network of relationships between *Non*

[30] It could be argued that Roesner defined his understanding of the term in general preface to the *Magnus Liber Organi* editions: 'a large number of clausulae, new settings of smaller segments of plainchant to replace or supplement portions of organa already composed' (Roesner 1993, p.xxxiv), but this definition clearly relates only to those so-called "substitute clausulae" and cannot be reconciled with or equated to his usage of the term elsewhere. By "large number of clausulae" he was referring only to "clausulae" as a genre distinct from the other material in the Magnus liber, but elsewhere he used the term "clausula" to refer to items embedded within the organa.

conturbetur and *Dum complerentur*. He was particularly interested in 'the end of the clausulae [on "tur"]' in both settings (Roesner 2001, p.252), where there is agreement in the cadential formulae. *Dum complerentur*'s "tur" can possibly be termed a clausula using Roesner's previous usage, as there is an independent clausula on "tur" in **F** (van der Werf 1989, p.108), but the same argument cannot be made for *Non conturbetur* because it exists only in two differing purum versions in the three main sources with no independent clausulae (van der Werf 1989, p.107). Additionally, Roesner discussed 'the clausula on "omnes" in *Repleti sunt omnes spiritu sancto*' (Roesner 2001, p.258) (which is in a similar clausula-less state to "tur" of *Non conturbetur*) and 'the clausula on "Viderunt" in the two-voice *Viderunt omnes, Notum fecit dominus*' (Roesner 2001, p.264) (which, although differing substantially between sources has no independent clausulae).

Perhaps, then, Roesner had expanded his definition of "clausula" slightly, to include swapping between organa settings without the go-between of independent clausulae? However, this cannot be the case as, on the contrary, he had in fact widened it further to describe any section regardless of style: in a footnote, Roesner cited 'the [purum] clausula on "Iudea"' from *Iudea et Iherusalem*, which he himself (in the same footnote nonetheless) had described as 'the same in all three sources' (Roesner 2001, p.264).

Against prevailing definitions of the clausula, then, where it is commonly held that a clausula must be in discant, there are alternate definitions, concurrent in academic literature, that would define a clausula as any section, regardless of its stylistic form, layout, or tenor movement. This terminological superposition has arisen due a lack of sharp definition of the term "clausula", where the term has become a useful bridge to cross the awkward void between lyrical, free-flowing organum on one side and rhythmic, strictly-defined motet on the other. This has then permitted conflicting definitions of the term to filter into this poorly-defined area. Nevertheless, despite the conflicting definitions, the term "clausula" is a Latin loanword. In borrowing the term, all of the definitions given so far have implicitly claimed knowledge of the medieval usage of the term, and thus an unbroken line of meaning from the thirteenth century to today. It may seem obvious to many therefore that, rather than continuing this historiographic discussion, we should instead discover the "true" usage of the term "clausula", so as to form the basis for our own definition.

2.7. The Anonymous IV passage

Unfortunately, the semantics of the term "clausula" in the thirteenth century are variable. We have already seen how Flotzinger (1980) had clearly stated that "clausula" had multiple meanings throughout the medieval period. A similar picture of multiplicity is given by Schmalzriedt (1974, p.1), and the *Lexicon Musicum Latinum Medii Aevi* (n.d.), which both divide the term into three concurrent and somewhat overlapping medieval definitions: a musical section, an ambitus, and a concluding passage. However, very few of these definitions concern the ND repertory directly, and they therefore often make little sense in the context of ND polyphony. For example, the Anonymous de la Fage treatise (twelfth century) uses the word “clausula” both as a term for “section” and “pause”,[31] but there is little that can be recognised in this usage for the ND repertory in the thirteenth century. The only clear corroboration of the term “clausula” with a writer discussing the ND repertory directly is Anonymous IV.

All the definitions explained throughout this chapter have therefore hinged on the modern understanding of the term “clausula” as used in Anonymous IV, first published by Coussemaker (1864–1876). Anonymous IV is the only theorist to deal with the ND clausula in any detail, and so discussions of the “true” meaning of the term within ND polyphony often revolve almost exclusively around the dissection of this writing. Unfortunately, the meaning of the text surrounding the term “clausula” in Anonymous IV is subject to much interpretation, due to its ‘clumsy if fussy Latin, which seems almost perverse in its search for precision’ (Roesner 2001, p.228). Therefore, different interpretations or translations of the text have aimed to untangle what it is that Anonymous IV meant exactly by *clausulae sive puncta*. These interpretations have had knock on effects on what modern scholars mean by the term as, understandably, these scholars have wished to replicate medieval usage. This is the same core issue that Planchart was attempting to explain in the quote that began this chapter: that the different interpretations of Anonymous IV have led to different understandings of the term “clausula”. Many of the historical definitions that have been surveyed so far have simply been glosses on the Anonymous IV text.

The key passage (see footnote on p.41) has been translated by, among others, Dittmer (1959a, p.36), Waite (1961, p.148), Yudkin (1985, p.39), van der Werf (1992, p.3), and Roesner (2001, pp.227–228). Each translation has attempted to understand the proximity of *clausula*, *puncta* and *optimus discantor*. For example, one could translate *puncta* as “sections” — as Dittmer (1959a) has done and as Yudkin (1980) has argued elsewhere — which would indicate that the term “clausula” is synonymous with “section”.

[31] The relevant passage is given by Handschin (1942, pp.24–25), and in translation by Sanders (1980, p.265). ‘But if by chance, in order to have a more beautiful and elegant discant as well as for the greater pleasure of the listeners, you should want to mix in some organal passages at the end of a period or section at the last or penultimate syllable of the text, that is permissible, even though the nature of the thing does not allow its inclusion’ (*Sed si forte in fine clausulae in ultima aut in penultima dictionis syllaba, ut discantus pulchrior et facetior habeatur et ab auscultantibus libentius audiatur, aliquos organi modulos volueris admiscere, licet facere, quamvis natura hoc non velit auferre*). ‘And indeed it may have a pause, which we call a *clausa* or *clausula*’ (*Nullam etenim pausationem quam nos clausam vel clausulam vocitamus*).

The linking word between the two (*sive*) has also caused consternation. It could be translated as "or" — which would indicate either that the two are the same or instead that the two in fact cannot be equated — or alternatively translated as "that is", indicating some form of clarification.[32] This must then be reconciled with Anonymous IV's further explanation that this was because Perotinus was better at discant than Leoninus, and the implication that discant is the key to understanding clausulae. One of the most recent opinions on this subject is Wegman (2015, p.719), who argues for the term's proximity to "section" rather than "discant":

> For Anonymous IV, the word clausula was synonymous with punctum. It had no precise definition, but referred generally to a stretch of notated music, often part of a larger piece, that was marked off either by a bar line extending from the top to bottom line of the staff […] or by a longa rest […] On one occasion Anonymous IV uses the term *clausa vel punctum* to denote a music example that contains two texted excerpts from a two-part motet, fused together as if they were one piece […] There is no compelling reason to assume, then, that the term clausula should have referred exclusively to untexted discant sections in organa dupla, or that Perotin's editorial efforts lay chiefly in the creation of such sections. The term could equally well have been Anonymous IV's preferred designation for motets, especially given that he avoided the word *motellus* itself.

It is important to bear in mind however, that this is only one interpretation of the passage. Wegman's analysis does agree with Planchart and others on the "true" meaning of Anonymous IV,[33] but it must also be accepted that Anonymous IV's writing is ambiguous and it may in fact be impossible to extract anything more than the series of interpretations we currently possess. Those that created the clausula-concept took the term "clausula" out of its context in Anonymous IV and co-opted it for its most important historical application: the transformation of organa into motet. In so doing, they forever changed its meaning in modern scholarship. However, Anonymous IV does not deal with motets, at least not directly, and was not thinking historically in terms of elevating the discant clausula as the precursor to the motet. As seen in this chapter, that was done at the turn of the twentieth century.

[32] For example, whereas other translations use "or", Roesner (2001, p.227) opts for "that is". See the further explanation for this decision given in Roesner (2018, p.834).

[33] As well as Schmalzriedt (1974, p.3) who considers the terms "clausula" and "punctus" as synonyms. Another important definition is that of Crocker (1966, p.76), who defined the clausula similarly to Reese (1940, p.298) and has therefore not been discussed here: 'The larger phrase was called a clausula or close (compare clause). Successive clausulas are often distinguished from one another by their style […] one of the most striking kinds of clausula […] moves in a regular rhythm'. If there is a "striking kind" of clausula that moves in rhythm, then it follows that there are other clausulae that do not.

2.8. The problem of the term "clausula"

The term "clausula" has become a metonym. Referring to one of the most interesting problems in the study of the ND repertory (the swapping of polyphony here termed the clausula-concept) has become synonymous with a particular, yet often unspecified, kind of polyphony (the clausula as a section). Regardless of how Anonymous IV (or any other medieval writer for that matter) understood the term, modern writers have used the term "clausula" for a variety of different purposes that cannot all be in agreement with medieval use. What can be inferred from this is that, far from modern scholarship being thought of as using a wrong definition against what we believe to be the "true" definition of the term, how the clausula was defined in the thirteenth century is in fact less relevant to modern study than how it is used in twentieth- and twenty-first-century academic literature. When scholars discuss the clausula, they are likely not concerned with the same aspects of the repertory that medieval writers such as Anonymous IV occupied themselves with.[34] The clausula that we see in the ND repertory is the clausula-concept, an image of what we may believe to be the "historical reality" (Wegman 2012).

Given this confused situation, it is vital to ensure that the communication of ideas within academic literature is clear and unambiguous. There has been scholarly disagreement about the meaning of the term "clausula" in the context of medieval writing on music, and so there is an accompanying need to create agreement on a definition of the term that can be used meaningfully within modern-day study. Different scholars have advocated for extremely varied definitions. This chapter has accounted for some of them, but there are many more that cannot be included here. However, they all lay some claim to the metonymic clausula, and usually to the modern clausula-concept that lies behind it. In summary, they can be divided broadly into five groups, here presented from the broadest to the narrowest, and expanding from Dittmer's (1959b) two perceived categories:

1. "Sectional": a clausula is simply a section of polyphony, typically supported by a generous reading of Anonymous IV (e.g. Planchart, Roesner, and Wegman).
2. "Musico-technical": a clausula is any discant section: this appears to be the most common usage today. If writers diverge from this definition, they typically mention it in a footnote.
3. "Historical": the clausula-concept is a process, cataloguing both the replaced and the replacing as a clausula. To speak of a clausula as an independent item makes little sense; it must be linked to what it replaces or is replaced by (e.g. Ludwig).
4. "Substitutional": a clausula is a discant section that is swapped, commonly found

[34] As mentioned on p.4, the aims of medieval theorists and modern writers rarely align (1993, p.11).

among the definition in mid-twentieth-century dictionaries and histories, this could be any discant section as long as it is substituted at some point.

5. "Codicological": a clausula is only those independent pieces preserved in separate clausulae fascicles (e.g. Flotzinger, Sanders, and Bradley).

These groups are not exhaustive but cover the majority of definitions in literature. Is it possible to reconcile these differing definitions? The first step a writer can take is to be aware of the divergence in the definition of the term "clausula" within scholarship. This is not simply a question of "old" definitions versus "new" definitions, but different groups of writers today are using almost opposing definitions of the term. Far be it for me to say which definition is necessarily "correct", but clarity in which definition a writer is using is to be appreciated. If it is acknowledged at the outset which definition is being used, then a reader can proceed knowing exactly what is being discussed and, more importantly, what is not. For example, it is clear that Büttner (2011) supports the "musico-technical" view:

> Klauseln sind Abschnitte des Organum, in denen sich die Tenorstimme rascher bewegt; man setzt sie im Regelfall dort ein, wo bereits der gregorianische Gesang, dessen Tonfolge ja im Tenor durchgeführt wird, ein längeres Melisma aufweist, dessen Zeitdauer so im Vergleich zum haltentonmäßigen Vortrag abgekürzt werden kann (Büttner 2011, p.12).
> *Clausulae are sections of the organum in which the tenor part moves more rapidly; as a rule they are used where the Gregorian chant, whose sequence of notes is carried out in the tenor, is already in a longer melisma, the duration of which can be shorter in comparison to a sustained-tone version.*

Another good example of definition is Bradley (2015) who appears to support the "codicological" view:

> I use the term discant here to refer specifically to a passage of discant preserved within the context of an organum. The term clausula is reserved for discant passages removed from their organal context and copied alongside other such clausulae in a separate section of the manuscript (Bradley 2015, p.154).

Crucially, these authors introduce their definitions at the outset of their discussions, and make it perfectly clear what is included within their parameters of the term "clausula" and what is not.

Unfortunately, it is impossible to attempt an adequate compromise between these definitions, not least because some are mutually exclusive but also that a compromise would only serve to dilute their individual purpose. Each definition is used as the basis upon which to make different arguments, whether that is the meaning of Anonymous

IV, the work of Perotinus, or an analysis of an early motet. Readers therefore need to be wary of the purpose of the term for, as we have seen, the term “clausula” very easily becomes a dangerously dynamic concept morphing to fill the conceptual space between other definitions of terms such as “discant” and “organum”. However, it is possible to evaluate which definition a writer generally uses, depending on their exact area of study. I do not claim these evaluations to be without exception or even particularly significant, but that they are general rules of thumb to have in mind when reading ND literature.

- If a writer means to make claims about the “true” definition of the term “clausula” as used by medieval writers, then they will use the “sectional” definition.
- Non-specialists and general histories will typically use the “musico-technical” definition: it is arguably the definition with the least inherent meaning, being almost entirely aligned with the term “discant”, as well as the most connected to the narrative strand of the clausula-concept.
- Writers that lean heavily on Ludwig or his cataloguers typically initially use the “historical” definition but, unless they are critical of this catalogue, will eventually turn it into the “substitutional” definition which can be found in most studies of clausulae.
- Finally, a writer primarily concerned with motet may use the “codicological” definition as a shorthand for “independent” or “substitute” clausula, as it is typically from these clausulae that the motet can be traced.

There is therefore no single or strict definition that can inclusively define what a clausula is without invalidating another definition, and any study must pick what is and what is not included in its use of the term if it is to be completely unambiguous. As such any definition of the term can only cover one or two of the five categories above.

2.9. Melodic formulae

The clausula is the first and largest item to be discussed within the context of musical reuse, but it is by no means the entire picture. The concept of musical reuse as something existing outside and independent of the clausula-concept in organum is a much less frequently discussed aspect of ND polyphony. Reuse and recomposition is of course well known in other genres of thirteenth-century music, most notably the motet, where the liberal reuse and reappropriation of refrains is key to the genre’s keen sense of intertextuality (Bradley et al 2019). Indeed, it may well be the case that the etymological root of the term “motet” derives not from a common polyphonic style but from the use of such intertextual monophonic refrains (Dammann 1959, p.349; .

Nonetheless, the ND repertory does include other, subtle forms of musical reuse within organum, as well as in conductus (Everist 2018, pp.199–213). While the clausula has dominated scholarly attention (by narrowing the broader view of musical reuse

Schmidt's (1931) first extracted formula from the ND repertory

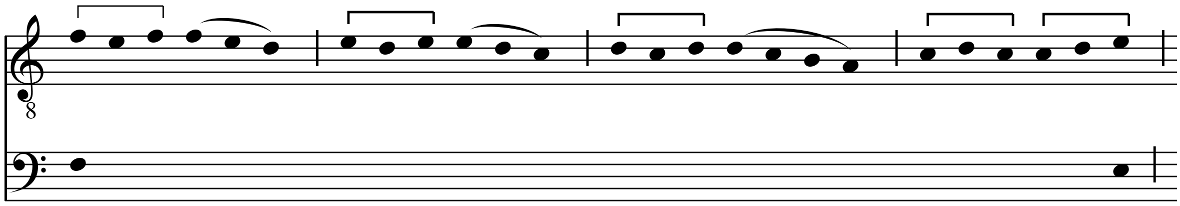

Fig. 2.3

Schmidt's (1931) second extracted formula from the ND repertory

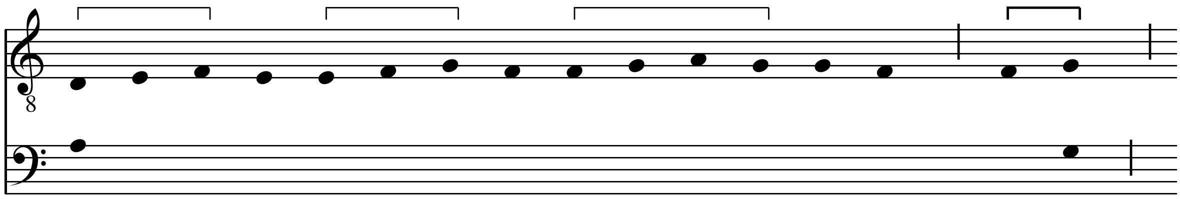

Fig. 2.4

within ND polyphony) research into a wider concept of intertextuality within the repertory has enjoyed a somewhat quieter academic trajectory. The clausula, which has often been framed as the historiographical linchpin that leads towards increasing compositional control and the emergence of the motet, contrasts with the looser, more fluid practices of musical reuse found in indirect concordances and melodic formulae. These practices instead suggest a culture rich in shared melodic patterns, stock phrases, and even improvisation (Treitler 2003b, p.71). Consequently, the broader study of reuse has largely avoided the intense scrutiny and stakes associated with clausula research.

The research progression consists mostly of coming to terms with the fact that musical reuse is indeed a key aspect of ND polyphony outside of the clausula-concept, and hence in the subsequent identification of a number of common melodic formulae, particularly at cadences. The earliest scholars in the field, such as Ludwig (1910), did not discuss reuse beyond the clausula-concept, but we can begin with Handschin (1928) who was the first to raise the question of what he termed "paraphrase" in medieval music. In so doing, he picked out a few relatively insignificant examples of internal repetition within the ND repertory, but it would be incorrect to say that musical reuse was indeed his main concern in this article. Erickson (1970, p.74) makes a good assessment of these first steps, in that Handschin's 'lengthy article finesses entirely questions of the phraseology of organum'. It was not until a few years after Handschin's study that Schmidt (1931, p.131) extracted the first true examples of melodic formulae from the ND repertory, and he did this through two examples of sequential cadential formulae that he found common between organal voices. The first (Fig. 2.3) he found in common between O2, O7, O9, M14, M17, M18, M28, M33, M46, M48, M55, and

M56, and the second (Fig. 2.4) in common between O1, O2, O3, O7, O12, O13, O14, O16, O18, O21, O23, O24, O28, O29, M20, M41, M46, M53, and M57.

Later scholarship saw a gradual proliferation of attention to such melodic formulae from within the repertory, based on a series of implicit criteria for differentiating between what constitutes a melodic formula, and what may simply be a chance encounter with similar melodic material. Karp (1966, pp.358–369) identified a formula in common between M3 and M11, Reckow (1967, p.82) found that a number of different cadences between settings share the same overall melody contour, Flotzinger (1969, pp.180–181) itemised 43 different common cadence figures, Erickson (1970, p.81) discovered that O10 and O11 'share a large number of duplum phrases', and Smith (1972) further identified numerous commonalities within both Alleluias and Graduals. However, this process of cataloguing formulae arguably reached its zenith in Tischler's (1988, p.lxiii–lxviii) comparative edition, where he identified 27 such formulae in duplum, and began to catalogue their use within the repertory. Schmidt's first formula is given number 23 in Tischler's tabulation, the second number 11, and some of Smith's are found in numbers 4 and 6.

Although Tischler's work does fulfil Erickson's (1970, p.75) desire to eventually create a catalogue of formulae, none of these writers went so far as to list the specific criteria that were used to identify such formulae. This issue was picked up by van der Werf (1991, p.952) who critiqued the lack of criteria, and described Tischler's results as 'uneven' and 'inconsistent' in this area. Tischler did not describe how he found these formulae, but simply listed them. It is likely that the lack of criteria in the volume is due to no specific criteria existing in Tischler's study, other than his own dead reckoning and sense that their presence is obvious simply by observing the notation. Admittedly, it is difficult also to see how suitable criteria for the identification of melodic formulae could be created: the same criticism could be said for any of the sets of formulae so far described, where all the authors could say with respect to formulae would be that "I know it when I see it". What may be a clear instance of a melodic formula to one author therefore may equally be dismissed on similar grounds by another.

The study of melodic formulae within the ND repertory has been more recently revitalised by an alternate perspective: rather than melodic formulae being a finite set of stock phrases that ND musicians selected from at will, formulae can be understood as carriers of a process reliant upon oral transmission and therefore subject to constant remaking and recomposition. This idea was already latent in Reckow's (1967, p.82) discussion of similar contours, but became more widely accepted with the advent of Treitler's (2003d) arguments for the extension of the Parry–Lord oral formulaic theory to Gregorian chant (Lord 1960), and Reckow's (1973, p.472) reframing of the "abbreviation" of the Magnus liber as more of a process of 'continuous modernisation' than intentional abbreviation.[35] Allison (2020, pp.25–48) traces the acceptance of this idea within the study of organum, and particularly influential here is Roesner's (2001) concept of "skeleton outlines". Roesner argued that the original "work" of organum purum comprised little more than 'prescriptions for fleshing out the musical substance of the duplum in performance', and 'points of departure' (Roesner 2001, pp.264–265). In essence, melodic formulae can be understood as the same flesh that transforms the skeleton outline into the fully worked-out-yet-variable versions transmitted in the sources.

In this more modern conception of musical reuse, investigation of formulae and indirect concordances is altered from a question of collating and cataloguing all possible instances of formulae to the more complex exploration of musical process occurring primarily within the mind of the singer. Here, what formulae and concordances are present within an organum matters less than how they emerge within the texture and how they are deployed by a musician. Roesner did not go as far as to discuss these formulae in any depth, besides a call for them to be 'systematically studied' (Roesner 2001, p.245). Since then, arguably the most thorough treatment of more subtle links within ND polyphony has been Roth-Burnette (2010), who demonstrated that undertaking this involved formula analysis for even a small group of settings is a formidable endeavour. As can be seen by the array of custom notation and sigla required for her diagrammatic analyses of only a small cluster of responsories (Roth-Burnette 2010, pp.254–255), the links within and between organa cannot be reduced to a series of simple one-to-one connections.

However, even with the involved analyses of those such as Roth-Burnette's, it is difficult to envisage what the final result is intended to be when applied to the repertory as a whole. This is because the internal model is much more challenging to conceptualise within this new phenomenological perspective of musical reuse in ND polyphony. In the previous model, Erickson (1970) imagined something akin to a large polyphonic concordancer, where the atoms of musical reuse could be extracted, processed, and presented in an exhaustive table. This is also what Tischler (1988) appears to have been attempting, albeit in a less rigorous manner. Roth-Burnette's model, however, consists of two parts: a thesaurus of melodic figures, and a grammar that informs a singer how to recompose (Roth-Burnette 2010, p.vii). This aligns broadly with Roesner's (2001) concept of fleshing out skeleton outlines. However, scholars have not yet considered the scale of the work that is set out for them in this model, and this raises more questions than answers: are we to expect to perform full-length studies,

[35] 'Overall, the reworking of the *Magnus liber organi* is far less for the sake of shortening than for the sake of continuous modernisation with regard to modal rhythm and periodicity' (*Insgesamt ist die Umarbeitung des* Magnus liber organi *weit weniger um einer Kürzung als um einer stetigen Modernisierung willen in Hinblick auf modale Rhythmik und Periodik).*

such as done by Roth-Burnette, for every complex of formulae that can be detected within the repertory? Such a work would have to culminate in some extremely large, multi-volume study that would need to outstrip all previous studies of the ND repertory both in scope as well as in detail. How do we set parameters on this study such that we know when our analysis of formulae is complete? More importantly, can this kind of subtle manuscript *mouvance* even be adequately discussed in a critical apparatus that 'hypostatizes to the point of unreadability' (Cerquiglini 1999, p.80)? How many notes and in what combination are notes required to construct a formula? How diligently must the melodic grammar be followed so as to be seen as significant?

Roth-Burnette is tentative on what is meant by "grammar" in this case. She adopts the term from Immel's (2001, p.166) conclusions on the Vatican organum treatise (Roth-Burnette 2010, p.15), and the word "grammar" is often found in her work within double quotes or as the compound adjective "grammar-like". In this definition of figures and grammar, it is understood that the reader agrees on Roth-Burnette's implicit and unwritten criteria for the identification of formulae. It is also understood that these formulae, once identified, can be recognised throughout the repertory by the same kind of duck test that is necessary for Tischler's list of formulae. However, what these criteria are is not made clear. Moreover, there may also be formulaic links that are transformed such that they cannot concisely be described by grammatical rules, thus the hesitancy in defining this grammar more precisely. If this hypothetical grammar within ND polyphony is anything like that of language, then there will likely be as many exceptions as examples that fit the rule. Furthermore, we must also be aware of the danger of apophenia here: if we begin with the assumption that there is such a musical grammar and that melodic formulae can be identified, then close reading will likely find exactly that by virtue of confirmation bias (Neuwirth and Rohrmeier 2016, p.176).

A similar critique along two lines is made by Allison (2020) in her study of purum: firstly that although Roth-Burnette finds melodic formulae practically everywhere in her analysis, they are short and 'not particularly characterful' (Allison 2020, p.45), and secondly that Roth-Burnette overstates her point about how conscious this process was when it could have been simply the result of musicians 'creating melodies like the ones they sang every day' (Allison 2020, p.46). Allison herself finds more examples of musical reuse in purum that better fit this "characterful" criterion, such as common patterns (pp.78–83), polyphonic building blocks (pp.138–140), and subtle variation over long periods (p.153). To my knowledge, this is the first instance of an author explaining how they are identifying formulae, and this is without doubt a positive development.

However the same issue remains where, unlike clausula concordance which is linear and one-to-one, the identification of melodic formulae relies on both author and reader agreeing on identical criteria for formulae and their possible transformations.

Even the Vatican organum treatise, which possibly transmits musical passages that could be understood as formulae, still requires a high degree of interpretation to put its formulae into practice. This extremely interpretative nature of formulae study has the potential to yield as many results as there are differing criteria. If we alter our criteria, we will likely end up with an entirely different set of formulae, and so this often tells us more about the author's reading than anything tangible or reproducible within the repertory. Unlike clausulae, which were discussed by contemporary music theorists and can be delineated in substitute fascicles, the identification of melodic formulae is almost entirely a product of what modern analysts have observed in the ND repertory. This is not to say that melodic formulae do not exist or cannot be detected, but that we cannot be certain that the criteria we are selecting for formulae accurately reflect the ways in which polyphony was created in the thirteenth century, and thus that the results created are significant.

2.10. Musical reuse in the computer

It is part of the discursive nature of arts and humanities subjects to treat exact definitions and interpretations with a high degree of scepticism, particularly in historical study. Postmodernism has long recognised that it is not possible to look back on the so-called "historical record" and extract anything approaching fact (Foucault 1972). What we do gain from looking at historical artefacts is instead 'how we have fashioned it in our image' (Wegman 2012, p.44). Musical reuse in the ND repertory has been subject to this repeated refashioning, and this chapter has shown how limited historical data has been interpreted and reinterpreted into evidence to support diverse theories, by generations of scholars, in wildly different ways, for various purposes. As mentioned on p.3, scholars working in the field today have access to largely the same set of data as those working a century ago. Divergence in opinion, modes of discussion, and ultimately, conclusions, are almost entirely due to changes of perspective — such as New Musicology or Material Philology — rather than any revolution in evidence gathering. It is this constant fluidity of new perspectives that lies at the root of the terminological differences that this chapter has catalogued.

Such differences are useful in scholarship precisely because they encourage the kinds of discourse that invite new perspectives and interpretations, each equally interesting and enlightening on their own terms. However, difficulties arise when the distinction between interpretation and historicity is lost, and the subjective impressions formed through academic study are conflated with the irretrievable nature of the past itself. The first is an interesting discourse that we might have in the here and now (our interpretation of the past), and the second perished entirely as soon as it ceased being the present (the "past itself"). Terminology that we use to discuss music history is entirely of this former, discursive type, and relates only to other historical discourse, never to the past itself. When writers use terms such as "clausula", their primary aim is to position themselves within that discourse, and stake their claim to an aspect of music history. For example, an author using the term "clausula" as a synonym for discant is, consciously or not, evoking ideals connected to the clausula's role in creating new forms within the ND repertory as well as its compositional aspects. Alternatively, an author using this same term to describe any section regardless of style is instead signalling their connection to a broader reading of Anonymous IV, and a more restrained opinion regarding the innovations of the discant clausula.

If this dissertation is to discuss the possibility of new computational directions for musical reuse, then we must first confront the fact that computers, for all their utility, do not work this way. Even the most sophisticated computer systems at our disposal do not engage in discourse in the way that we as humans do (Ramsay 2011, p.8). Computers process data, only the data that is given to them, and then only in the forms that they support. Computers alone are not capable of altering their perspective in the same way that we approach musical and historical discourse: they must be told to do so as part of their model, at which point the model becomes not a true engagement with discourse but a perspective-altering computer program. In programming computers we tell them what to do and how to do it, and as such we imbue our computer programs with our own perspectives and biases (Marino 2020, p. 35).

As a result, creating new computational directions for studying musical reuse in the ND repertory is not an action that can exist outside of or detached from the study of musical reuse more generally. Rather, our own perspectives are automated — and therefore amplified — by the speed of the computer. Any of the perspectives on musical reuse that this chapter has surveyed could well be encoded and operated by a computer, some more easily than others. As stated in "1.2. Methodologies", the aim of this study is to fit the computer to the questions we are asking, not to ask what it might be possible for the computer to do. However, study that has not used the computer exists within discourses that frame their investigations in forms that computers fundamentally cannot accept.

For example, it would be possible to harness the power and speed of a computer to perform a study such as Roth-Burnette's (2010) at the scale and scope required to encompass the entire repertory. It would first be necessary to encode precisely the criteria for the detection of melodic formulae, and then run those criteria against every pair of phrases within the entire repertory. This would require a complete reframing of the kind of work Roth-Burnette achieved. The results would be precisely as I have described: we would gain thousands of pages of possible links between settings.

However, we must also ask whether such an outcome is in fact desirable. We already know how Roth-Burnette conceives of organum in terms of a musical thesaurus and melodic "grammar". Her interpretation of musical reuse is what is useful to the academic discourse, not the exact outcomes. If we change the criteria, perhaps after Allison's (2020) more "characterful" suggestions, we will end up with different results, but this will not tell us anything more than what we already know about both writers' work from their own words.

A computational replication of others' work, then, is not a useful direction, and may serve only to cheapen the arguments originally made. More useful is a direction that uses the particular features of the computer to its advantage. The tools that we use shape the ways in which we approach and process our materials. Simon (1996, p.12) characterises this as a dialectic between inner and outer systems, where although the results gained show something of the outer system (the environment we are wanting to study), the use of tools themselves begins to show us more about the structure of the inner system (the internal structure of the tool we are using). The ways in which musical reuse is conceptualised and discussed are tools, and this chapter has shown that these tools tell us about themselves as well as about musical reuse itself. What is common between all the tools so far introduced is that they rely on certain criteria: definitions of "clausula" rely upon sectional, musico-technical, historical, substitutional, or codicological criteria, and discussions of melodic formulae rely upon the authors' written or unwritten rules for detection.

Box & Draper (1987, pp.4–10) discuss this as a general issue throughout scientific enquiry that, given identical research questions, different experimenters will design different experiments. This chapter has seen how musical reuse has been conceptualised within various models such as the clausula-concept, sectionalisation, and melodic formulae. Box & Draper's proposed solution is iterative, based on Occam's razor: that experimenters should begin with simple, continuous models before alighting upon any categorical responses (aka criteria). Thinking about musical reuse in terms of criteria limits the scope of reuse into a series of categorical statements, "binning" musical reuse into a finite number of types: e.g. is this an example of musical reuse? what category of reuse is this? what kind of connection links the component parts? My own interpretation is that this precipitation of musical reuse into a set of criteria has the potential to create fairly limited or even misleading results, based more upon the inner rather than the outer system. Simpler, continuous models assessing the level of musical reuse of any type have not yet been attempted.

Humans, of course, are not very good at dealing with continuous data. Not only do we prefer the certainty of a categorical outcome, we also lack the mental arithmetic required to balance numerous quantitative measures and express them by degrees.

Computers are the exact opposite: they excel at performing this arithmetic as well as storing huge sums of data, but they lack the interpretative power necessary to create categorical statements. This study can make an improvement to the issues surrounding musical reuse in the ND repertory by positioning itself at this human-computer interaction. By stripping back the layers of categorical decision, we can create continuous models that do not rely on criteria, and instead consider musical reuse as a single continuum. From this outcome, the resulting distribution can be analysed. If necessary, criteria can then be extracted from this distribution and the model iterated into something a little more complex and categorical. However, before performing these formative experiments, we must first decide on the data, data structures, models, and methodologies that will allow this to be possible, which is what the next chapter will discuss in detail.

Chapter 3
Data models

A key prerequisite to identifying the effectiveness of new digital methodologies for extending the analysis of musical reuse in ND notation — that is, beyond the clausula and the development of criteria for musical formulae — is the creation of a reliable computer encoding for ND notation that would allow us to extract patterns of musical reuse. To accomplish this, we can use the fundamentals decided upon in "1.4. Another look" to discern exactly what we know about ND notation and therefore how best to approach modelling it.

Digital encodings of music, even those that use the same notation system as the source,[1] are equally as editorial as those we might more readily term an edition printed on paper. The remediation of content from one medium to another necessarily requires some level of (re-)interpretation and (re-)construction, whether that is a simple symbolic mapping from pen strokes to font, or a more involved transformation such as halving note values. When discussing original sources, there is no substitute for the sources themselves (Bent 1994), where even the original spacing can be seen as a vital quality of the manuscript's materiality (Bent 2013, p.249). Cerquiglini (1999, p.79) has critiqued such diplomatic transcriptions that aim to reproduce the original sources as closely as possible by maintaining that 'every edition is a theory', even diplomatic transcriptions. As long as we are transforming content between media, we are interpreting that content at some level.

The same is true for computer encodings of music: they are models and theories of how that music operates, and therefore require certain assumptions to be made concerning the nature and range of behaviours of that item. Internally, a data model is precisely that, a "model" of how a real object works, not the object itself or an entirely transparent copy. The discovery of the required range of a model is known as domain analysis, where the domain defines the conceptual level that the model will operate upon (Smiraglia 2015).

This links directly to McLuhan's famous maxim 'the medium is the message' (McLuhan 1964, p.7), where, rather than making the simplistic point that the message itself is inconsequential, McLuhan was making the point that the impact of any medium is insidious precisely because it is hiding in plain sight. What we consider the "content" is just 'the juicy piece of meat carried by the burglar to distract the watchdog of the

[1] That is, if it is at all possible to discuss any early notation as part of a unified system. See Bent (2013, p.245)

mind' (McLuhan 1964, p.18). And what of the medium? That is 'any extension of ourselves' (McLuhan 1964, p.7). The tools that we use (the medium) to approach a problem (the message) affects not just the content itself, but in fact frames the content in a particular way and alters the ways in which it can be transmitted, transformed, and analysed. McLuhan makes a point of print being a medium for writing, and writing a medium for speech (McLuhan 1964, p.8). What we can do with the underlying content (speech) differs depending on the medium in which it is encoded. Similarly, an image of a page of text contains the same underlying "content" as a symbolic representation of that same writing in a text file, but the framing media ascribe different contexts to it and dictate what can and cannot be done with it. For example, the image contains layout and margin information, but the text file can be edited and copied into new layouts with ease. Although the underlying message is the same, the media that transmit the message inhabit different domains: speech in sound, writing on the page, text in a digital file.

For this study, the initial problem domain is graphical (musical notation on the written page) and from this I am trying to extract an analytical domain (patterns of musical reuse). This is not a simple conversion, and a large part of this study's methodology will be occupied with discovering efficient ways with which to effect this conversion process. A common way of doing this is through transcription, i.e. converting the notation into a form approaching CWMN. However, this is not an ideal methodology as, viewed through McLuhan's theory of media, when we transcribe we are altering the medium by which the notation is represented. Not only do any transcriptions of ND notation into CWMN alter the content of ND notation itself from being partially descriptive — such as unmeasured or partially specified rhythm, as well as ambiguous alignment — to being fully prescriptive, but they then frame the content as being just like any other CWMN, thereby altering the ways in which it can be further processed. The editor is involved in a process of converting the notation from the graphical domain (the ND notation) into a symbolic domain (their interpretation of the semantic content), and then back into another graphical domain (CWMN). This will not be the approach taken here.

To the best of my knowledge, there have been two previous attempts to encode ND notation without first transcribing it into CWMN: Erickson (1970) and Stutter (2020). Each took a different symbolic approach to encoding ND notation. Erickson (1970) encoded the notation in a tabular format, with the alignment of specific pitches between voices accomplished by the alignment of rows and columns in the table. Conversely, in Stutter (2020), I used a hierarchical format to represent groups of notes occurring within the same time frame, periodically dividing the structure to represent a broad level of alignment. However, although I achieved a more flexible encoding scheme than Erickson (1970) for areas where the alignment was unclear, dividing the entire structure into discrete periods caused issues when voices began overlapping. This was not a problem for Erickson (1970) as he was dealing only with unmeasured two-voice purum, but the flexible representation of both measured and unmeasured music in two, three, and four voices is as yet an unsolved problem.

One important and easily overlooked aspect of these previous attempts is that they began from a symbolic domain: the authors first interpreted the graphical domain before encoding it into the computer. However, the interpretation does not always fit with the model that has been created for it, leading to undesirable results and a very clear example of the medium transforming the message. This chapter will investigate the main models of music encoding from a conceptual level with a view to selecting the most apt model to encode ND polyphony for analysis beginning not at the symbolic, but at the graphical domain of the source material. I will then survey common methodologies used to find answers to the research questions outlined in "1.1. Research questions", and how they may be adapted to investigate musical reuse.

3.1. Ontology, epistemology, domain, representation

However, before a framing for the representation of ND notation can be created, it is vital that the philosophical underpinnings of domain and representation, namely ontology and epistemology, are considered. Despite multiple encoding systems being based on taxonomies of domain — such as the difference between graphical, gestural, and symbolic notation — a critical evaluation of what domains exist in music and how they should be related has not been attempted. Therefore, my first task will be to survey the literature that defines musical domain and calculate how the domains of this study will relate.

A theory of music ontology is one that is concerned with the existence of musical works and the relations between them (Kania 2023). An epistemology of music, on the other hand, is 'concerned with how it is that we know what we know' (Cross and Tolbert 2021). When it comes to music ontology, claiming to encode "music" in the computer using a data model is an audacious stance. What exactly is meant by "music" here, and does this object exist? For example, Matheson and Caplan (2011, p.47), in considering the example of Beethoven's *Hammerklavier* sonata with respect to ontology, must first assume that the object exists before they can consider any ontological questions surrounding it. Indeed, the first question asked when considering musical works is: 'are there any?' (Kania 2012, p.198)

This is an especially difficult question to answer in music, as what we consider "music" encompasses all manner of manifestations. To ask "what is music?" is to open a much broader premise: when we talk of music we are speaking of something that could take many forms. Our music could be auditory, visual (notation), or both, or something else entirely. It goes without saying that I will not seek to answer this question here, but instead highlight the difficulty and its ramifications for epistemology. Cook (2002, p.78) admits that 'you can easily find yourself asking, without any clear sense of what the answer might be: is this theory about acoustic events or perceptions, about notational traces or ideal content?' In taking the "audacious" step to encode music, we are saying that it is possible to: a) distil the consistent parts of a music-object from out of one or more of its many manifestations (ontology), b) say what we know about that object (epistemology), either by stating facts about what we consider the music to be, or by drawing a boundary around the finite set of its possibilities (Pace 2009), c) represent and encode it in a computer. This stance is necessary, as a computer is a machine that requires state to function.

Commonly used terms within music encoding that address this stance are "domain" and "representation". Domain is related to the mathematical concept of domain — being the range of acceptable inputs to a function — for example a pitch domain could consider relative pitch as the letters A–G and its respective octave. This input would not, however, be acceptable to a pitch domain that was accepting sharps, flats, and quarter tones.

Representation, on the other hand, is concerned with the means of encoding, and how the inputs of a domain are represented in the computer. Crucially, representation is not what to encode, but how to encode. An example of this would be the decision to represent a relative pitch as its letter and octave (perhaps including accidental) or simply as a number of semitone steps from middle C. Each represents the same domain (pitch) but with different power: the former allows for differentiation between A-sharp and B-flat, whereas the latter can be more easily interpreted into frequencies for computer performance. Following McLuhan, the medium of encoding, both through its domain and its representation, alters and places limits upon the message of the encoding. If the domain does not support an input, then that input must be left out or altered to fit. Similarly, if the representation is not fit for purpose then an important distinction, such as between A-sharp and B-flat, may be lost.

3.2. Attempts to define domain

It is impossible to create a music representation system that can represent all music, as it must be able to encompass all possible domains of music and organise them into a taxonomy (see Gödel's incompleteness theorem). Music representation systems must therefore limit themselves to a specific model of music through their chosen domains and representations. However, if music representations do not consider the problem of domain and therefore their inherent limits, they risk confusing or conflating aspects of one domain (such as a graphical domain) with another (such as a symbolic domain). As we increase the number of domains across which a music-object is encoded — graphical, symbolic, analytical, bibliographical, etc. — it begins to indicate that the object in question does not exist, or at least is dispersed among entities such as notation and performance. If the music-object — such as the *Hammerklavier* sonata — does exist, then considering its existence through a large number of domains implies that it cannot be fully defined through any combination of them and is therefore a Platonic ideal. Rather than being a "real" object, the music-object exists through its artefacts and manifestations: these manifestations can tell us something about the music but they are not the "music itself".

Domain is therefore used as the constituent for finding common ground between the manifestations of music, based on McLuhan's theory of media: if it is possible to represent multiple domains through multiple media within a single encoding, then extracting what is common between them will be the essence of the music contained therein. Put another way, attempts that claim to encode Beethoven's Op. 106 must posit that the *Hammerklavier* sonata is the sum of all the commonalities between all representations that suppose to be of it, and that there is nothing about the *Hammerklavier* sonata that cannot be represented in an encoding. After all, it is impossible to encode the *Hammerklavier* sonata directly, and so it must be done through its related manifestations such as scores, performances, and metadata. I do not wish to support or deny this view, only to highlight that music encodings must make an ontological decision as to what they are encoding in the first place, in what domain, and for what purpose. Consequently, any encodings that purport to transparently represent the "music itself" should be treated with a high degree of scepticism.

Despite this, there is little agreement in the literature as to what the domains of music should be as they pertain to music representation systems. Ontology is a common theme around which discussions of domain are based; for example, Steyn (2013, p.5, emphasis original) divided music into 'a core which is *music itself*, and around which all other themes revolve'.

Babbitt (1965) is commonly cited as the first to divide musical experience into domains (although it could be argued that the taxonomy of music is at least as old as Boethius' sixth-century *De institutione musica* and his classification of music into *musica mundana*, *musica humana*, and *musica instrumentalis*). Babbitt himself cited Kassler as the progenitor of the concept of musical domain, although he did not say where.[2] There have since been numerous taxonomies of musical domain suggested. The following is a list of musical domains found in academic literature. The list makes no claim to be exhaustive, but is intended to demonstrate the variety of domains

[2] The terminology appears in Kassler (1963), but not in the form that Babbitt quoted.

Summary of domains listed in literature

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Babbitt (1965)	×	×		×				
Maxwell (1981)	×	×		×				
Huron (1992)	×	×	×	×	×		×	×
Leman (1993)	×	×	×					
Sloan (1993)		×	×	×	×			
Selfridge-Field (1997)	×	×	×	×	×			
Haus and Longari (2005)	×	×	×	×	×	×		
Steyn (2013)			×	×	×	×		×
Roland et al (2014)		×	×	×	×	×		
Baratè et al (2016)	×	×	×	×	×	×		

Tbl. 3.1

available in representation systems. There are slight differences in the definitions between authors but, for the purposes here, I believe these not to be significant. In each case, I have taken the earliest citation:

a) **Audial:** This relates to sound waves: the ephemera of music or audio recordings. Babbitt (1965) calls this 'auditory', Maxwell (1981) 'physical', Huron (1992) simply 'sound', Leman (1993) 'acoustic', and Selfridge-Field (1997) 'phonological'. Haus and Longari (2005) and Baratè et al (2016) term it 'audio'.
b) **Performative:** This relates to the ways in which performers create music. Babbitt (1965) calls this 'acoustic', Huron (1992), Haus and Longari (2005), and Baratè et al (2016) 'performance'. Leman (1993) calls it 'subsymbolic'. Sloan (1993), Selfridge-Field (1997), and Roland et al (2014) term it 'gestural'.
c) **Logical:** This relates to a semantic interpretation of music, typically an abstraction from its graphical domain. Huron (1992) calls it 'common musical notation', Leman (1993), Sloan (1993), and Roland et al (2014) term it 'symbolic', Selfridge-Field (1997) 'the semantic context of musical perception and understanding', Haus and Longari (2005) 'music logic', and Baratè et al (2016) 'logic'. Steyn (2013) terms it 'concept'.
d) **Graphical:** This relates to the notation of music on the written page. Babbitt (1965) calls it 'graphemic', Huron (1992) 'visual notation', Sloan (1993) 'visual'. Haus and Longari (2005) and Baratè et al (2016) term it 'notational'. Steyn (2013) terms it 'written format'.
e) **Analytical:** This relates to the structural or theoretical analyses of music in another domain. Haus and Longari (2005) and Baratè et al (2016) term it

'structural'. Steyn (2013) terms it 'theory'.

f) **Bibliographical:** This relates to all forms of structured or unstructured metadata. Haus and Longari (2005) and Baratè et al (2016) call it 'general'. Steyn (2013) terms it 'language'.

g) **Meta-scores:** Huron (1992) adds to the graphical domain the domain of meta-scores, encompassing notational representations that set out the processes by which a score may be created. He gives as examples 'Xenakis-like tendency masks, tables of conditional probabilities associated with information–theoretic analyses, self-similar or recursive processes, transformational–generative grammars, etc.' (Huron 1992, p.15).

h) **Digital:** Huron (1992, p.12) also adds as a domain the intermediate representation of 'sound synthesis information', being one step removed from the audial domain and including all models of physical sound, such as 'algorithms and note-lists'. Steyn (2013, p.9) includes a similar meta-domain, that of 'markup expressions' relating to the markup for the encoding itself, such as *Extensible Markup Language* (XML).

Tbl. 3.1 summarises the kinds of domains defined in selected literature, and aligns them according to the eight categories outlined. What can be drawn from this table is that what constitutes "music" is different in each representation's case.[3] Scholars have found it difficult to create exhaustive lists of musical domain, as such a list is likely impossible to create. The search for all domains of music is open ended, and may ultimately be fruitless. As soon as you believe that you have captured all domains, there is always another domain that you have not considered, as what constitutes "music" in this case is not well defined. However, the best an author can do in these circumstances is to limit the scope of their schema by limiting the number of domains included in their representation system. Most of the domains cited above, for example, are considered almost exclusively within the context of Western classical music and may not be useful for this study. Attempting to encode ND polyphony within a CWMN context would cause the message to be altered by its medium (the representation system and its constituent domains) but unfortunately there is currently no better alternative to this scenario.

There is therefore no way for this study to avoid this problem: the domains and representation that I will ultimately decide upon for the ND encoding system will irreversibly alter its message, limiting the "what" and "how" of the encoding. The best that can be hoped for, then, is to be pragmatic about what can be achieved within the boundaries of this study, and recognise that even if the encoding that I create is suitable

[3] Although it is understandable that earlier lists do not include domains reflecting more recent technologies.

Wiggins' (2009b) representation of Babbitt's (1965) three domains

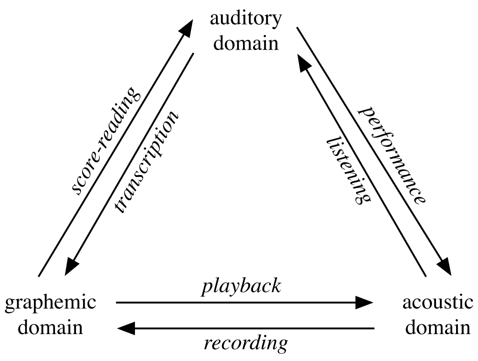

Fig. 3.1

From Haus and Longari (2005)

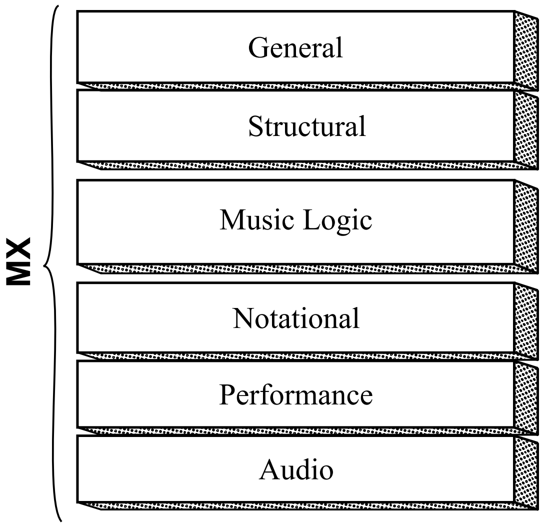

Fig. 3.2

for the domains of this study, it will not necessarily be suitable for another such study.

The aim of this study is to create an analytical domain from the graphical domain, and therefore it is useful also to consider the interactions possible between these two domains and others within the representation, as they may be implicated in the conversion process. For example, how is the notation (the graphical) interpreted in performance? Or how does our on-paper analysis of the score relate to the logical underpinnings of crotchets and quavers? Encoding schemata are often opinionated on these relationships, too, and so the possibilities of representational structure must also be surveyed. A similar unfocused picture can be seen here, and is best demonstrated by the diagrams authors use to organise their concepts of "domain".

3.3. Representational structure

Figs. 3.1–3.5 showcase some ways in which authors have related their domains to one another. In the most basic structure, Fig. 3.1 considers Babbitt's (1965) three domains each relating one to another: this is a simple structure to create when only considering three domains, but understandably this becomes more complex when more domains are added. A fourth domain such as analysis would have to be related to the auditory domain (e.g. an analysis of performance), acoustic domain (e.g. an analysis of a waveform such as a sound file), as well as the graphemic (e.g. a more traditional score-based analysis). This describes a complete graph, where the number of edges in a graph with n nodes is ${}^{n}C_2$ or:

$$\text{edges} = \frac{n\,(n-1)}{2}$$

From Huron (1992)

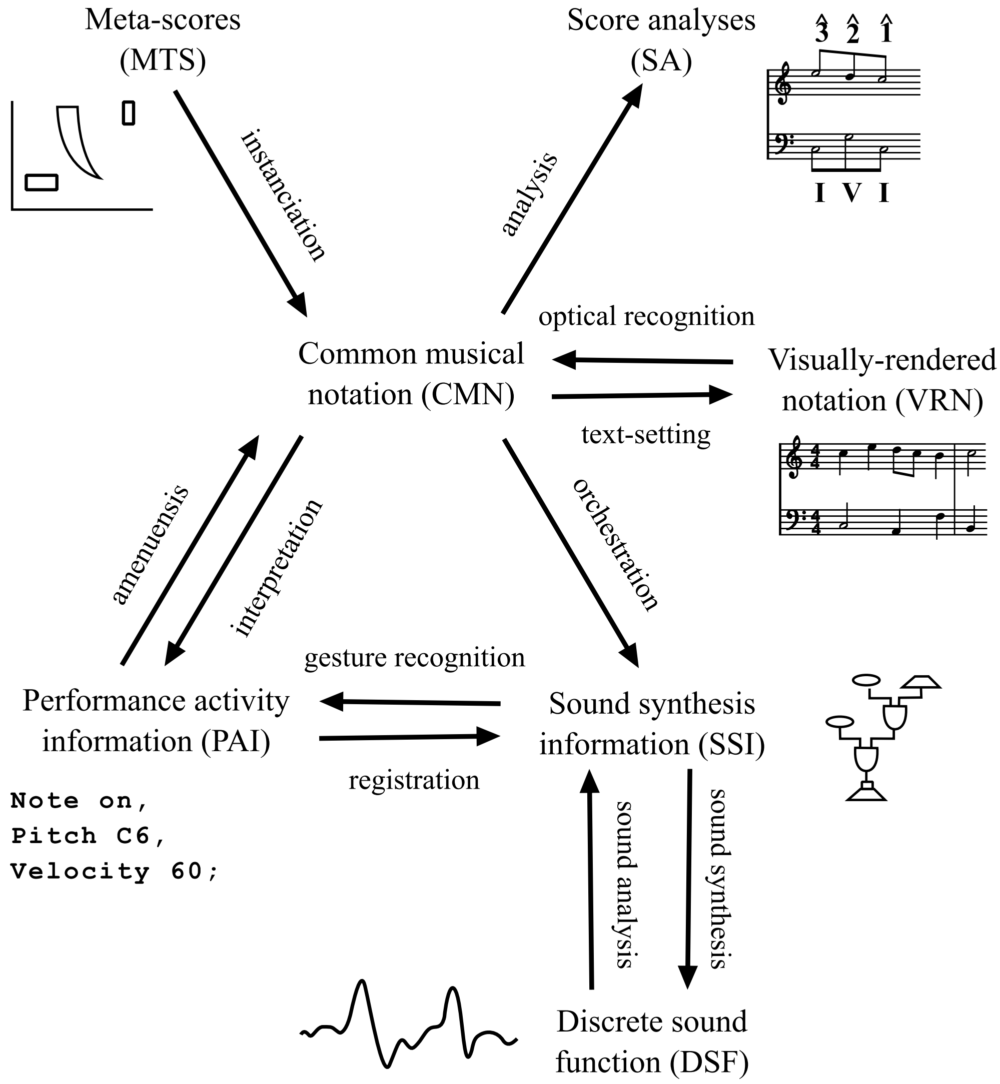

Fig. 3.3

As n increases, the number of edges in the graph increases quadratically. The three domains in the figure only require three edges, but a fourth analytical domain bumps the number of edges up to six. If we included all eight domains surveyed in the previous section, then it would require as many as 28 relationships to be described (or double this if we consider each direction as an independent relationship). Most encoding schemata would find this cumbersome, and so they introduce hierarchies into their relationships. One simple way can be seen in Fig. 3.2, where Haus and Longari (2005) stack their domains in an order: at the lowest level is the audial domain, and at the highest a bibliographical (general) domain. Each subsequent higher level is ‘an abstraction step’ of the lower, whereas the audial domain is the ‘layer closer to “what we hear”’ (Haus and Longari 2005, p.73). In Fig. 3.3, Huron (1992) goes further, and relates all domains through a central hub of so-called “Common musical notation”, which in this case is the same concept as what I have so far been terming CWMN. The two exceptions to this hub and spoke model are: digital sound synthesis information, which according to

Huron is a product of the "orchestration" of CWMN and the "registration" of performance information, and the audial itself (here called the "discrete sound function") which is solely a product of the sound synthesis information.

In both cases, most domains are not connected directly but information must often flow through another domain to reach its destination. For example, in Fig. 3.1 the transformation of the graphical to the audial domain is described simply as "playback". However, in Fig. 3.2 such a connection is not possible: data from the graphical (notational) domain must first flow through its lower layer (performance) before reaching the audial (audio) domain. Implicitly, a claim is made that the path from the graphical to the audial domain is equal to the concatenation of the path of the graphical to the performative and the performative to the audial:

$$(\text{graphical} \rightarrow \text{audial}) = (\text{graphical} \rightarrow \text{performative}) + (\text{performative} \rightarrow \text{audial})$$

Such a claim is contentious because it is no less than a theory of music based on an unfounded belief: that music must flow through a series of processes within these discrete layers to be transformed from one form to another.

Even more fundamentally contentious is Huron's (1992) implicit claim that all music must flow through CWMN (Fig. 3.3): for example it is impossible, in Huron's view, that a meta-score could be performed in an encoding without first being "instanciated" into CWMN. Before making such a claim, the structure needs to be sure that there do not exist meta-scores that do not require this step. On the contrary, many graphic scores for performance exist precisely because they cannot be represented in CWMN due to aspects such as aleatoric elements that require direct connection between score and performance. There is no way, for example, to "translate" a score such as Franco Donatoni's *Babaì* (1963) into CWMN:[4] it requires performance directly from its original graphic score. The same case could be made for pre-CWMN such as ND notation as a form of meta-score which, after McLuhan, cannot be losslessly interpreted into CWMN. Therefore, requiring a separate instantiation step limits the possibilities for musical process. This argument against domain-structural-rigidity can be applied in other cases, such as in Lindsay and Kriechbaum (1999) (Fig. 3.4) where the domain relationship is more complex, but yet is the only possibility allowed for such relationships in this representational structure. In this case, each layer must move through the "architectural" layer to interface with other domains.

There is no consensus nor sound academic framework to support the theory that the domains of music have such a set of hard-coded relationships between them. Although many relationships between domains may be common, such as the conversion

[4]Reproduced in Cage (1969).

From Lindsay and Kriechbaum (1999)

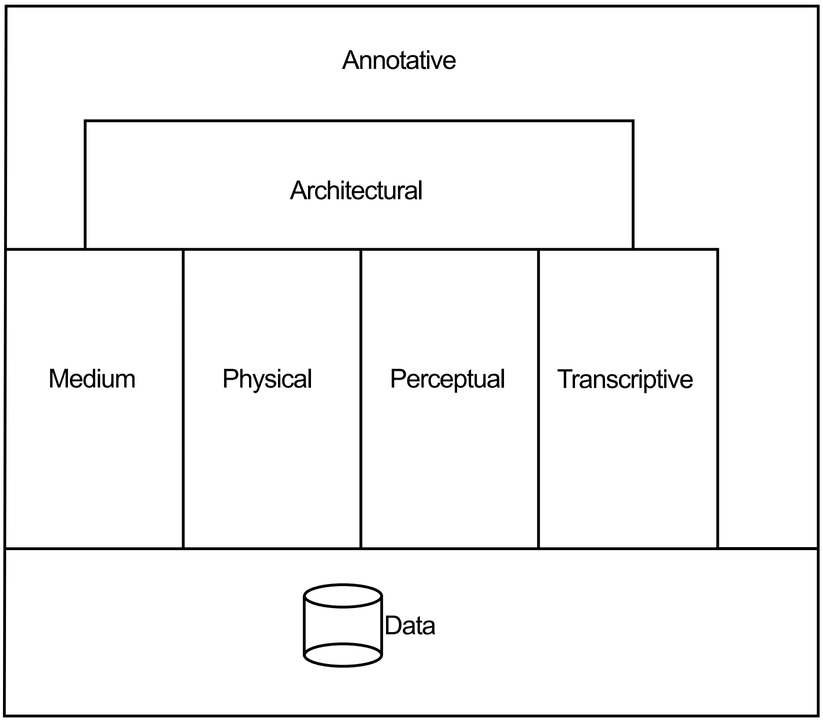

Fig. 3.4

Adapted for clarity from Baratè et al (2019)

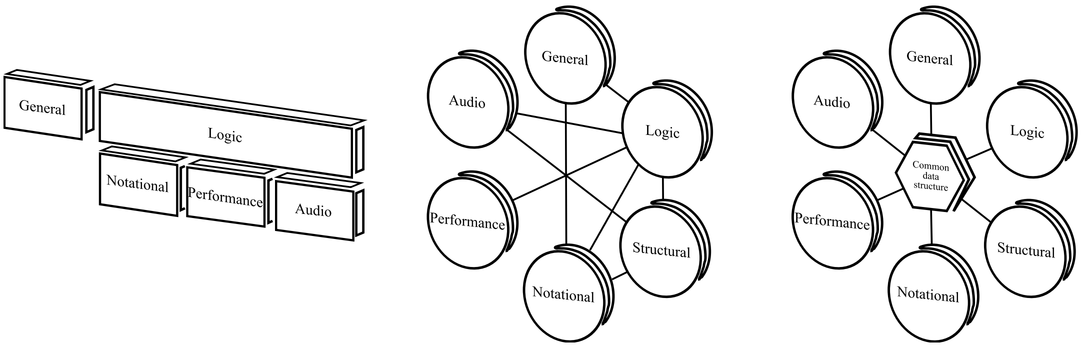

Fig. 3.5

from the graphical to the symbolic domain in reading music either manually or via OMR, these are not the only means by which the domains of music can be related. I have already argued that the number of domains available to us is not exhaustive, therefore if we were to add a new domain we would also need a new set of relationships between the domains. To my knowledge, the only attempt to create a set of more fluid domain relationships in alignment with this is Baratè et al (2019) who, although limiting themselves to the six domains first outlined in Baratè et al (2016), maintain that the relationships between them can take many different shapes. They give three examples of possible models (Fig. 3.5), not just the hierarchical and hub and spoke models, but also a more free-form graph model (middle example).

Two conclusions can be drawn from this survey. First, the number, size, and purpose of musical domains are not decided. Although there is a general trend of more domains being added in more recent literature — from three in Babbitt (1965) to six in Baratè et al (2016) — exactly what these domains are and how they should operate varies depending on the encoding standard being discussed. Second, the way in which these domains are related to each other is also undecided and each set of relationships communicates an unwritten conceptual theory of music based upon the author's preconceptions of how music works. None of these models are proven, and they often form a rigid structure based upon the processes most commonly seen in Western classical music. These processes cannot be transferred directly onto ND notation, and so this study must use a set of domains and representational structure tailored specifically not only for ND notation but also for this study's particular aims, as I will not be considering domains such as performance directly.

3.4. A framing for ND polyphony

For the purposes of this study, it is assumed that a data model that will be suitable for encoding ND notation with the purpose of analysing for musical reuse can be conceived within this same framework of domain and representational structure. Given the previous sections' findings that these concepts are not well defined, the study must begin from first principles with the knowledge that the requirements may not be catered for in pre-existing encoding schemata. It therefore needs to identify exactly what domains are involved and how these domains' relationships should be theorised. "1.4. Another look" was a conceptual data model of sorts, and this section will move towards creating a logical data model of the same. The physical data model will be created in "5.1.2. Database design".

At its root, this study is interested in the relationship between the notation of the ND repertory and the structural domain of musical reuse, such as in clausulae and individual melodic formulae. The main research question could therefore be rephrased as an interrogation of the link between a graphical and analytical domain. There is also a small bibliographical domain: "1.4. Another look" began from the premise that the music contained within the sources is divided into multiple settings, and that the sources themselves have various metadata associated with their continued use and storage in their host institutions and online. This study is not interested in domains such as performance and audio, and so my encoding system need not have the capabilities to encode this data.

How then are the graphical and analytical domains related in this case, and must other domains be introduced in order to facilitate their relationships? To determine this, I will consider the parameters of each domain for some representative data, formulate this into a data model, and then theorise how these domains are related. Simsion and Witt (2005, pp.10–15) outline nine criteria for a good data model: completeness, nonredundancy, enforcement of business rules, data reusability, stability and flexibility, elegance, communication, integration, and conflicting objectives. The final two criteria are not relevant here as this is an isolated new model, but the created model will be evaluated using the remaining seven criteria.

3.4.1. Graphical domain

"1.4. Another look" identified the following elements within ND notation: notes (on or off the stave), plicae, staff lines, divisiones, lyrics, clefs (of numerous types), accidentals (sharps as well as flats), and ligatures (square and currentes). Ligatures are a hierarchical structure, containing multiple notes. Within a graphical domain, each item may be attached to an *xy* coordinate on the stave, but may also contain some other information concerning its size and orientation, as well as its type. Using information from the first chapter, the required data for the graphical domain can be identified as:

- **Notes** (a single point):
  - Note ID number
  - Note *xy* coordinate
  - Note *plica* direction
- **Staff lines** (a line):
  - Staff line ID number
  - Staff line *xy* coordinate 1
  - Staff line *xy* coordinate 2
- **Divisiones** (a line):
  - *Divisione* ID number
  - *Divisione xy* coordinate 1
  - *Divisione xy* coordinate 2
- **Lyrics** (boxed text):
  - Lyric ID number
  - Lyric *xy* coordinate
  - Lyric width and height
  - Lyric rotation
  - Lyric text
- **Clefs** (boxed type):
  - Clef ID number
  - Clef *xy* coordinate
  - Clef width and height
  - Clef rotation

  - Clef type (C, D or F)
- **Accidentals** (boxed type):
  - Accidental ID number
  - Accidental *xy* coordinate
  - Accidental width and height
  - Accidental rotation
  - Accidental type (sharp, flat, or natural)
- **Ligatures** (list):
  - Ligature ID number
  - Ligature's list of contained notes (length > 1)

Purposefully omitted here are any qualities of notes, such as size or duration. For example, some longas are distinguished between simplex and duplex by the horizontal length of their note heads. However, ascertaining the point at which a longa should be interpreted as duplex rather than simplex is a difficult and contextual process not necessary for the data collection of this study.[5] In alignment with the needs of this study and according to the criteria set out by Simsion and Witt (2005), this model supports all the necessary data in this domain for ND polyphony in this study. Needless to say, a different study looking at different parameters would require a different set of data: for example, how this data is distributed over systems, or marginalia. This study is not interested in questions relating to such other parameters, so only some elements of the notational data is required, although the model does not preclude extension to further attributes. There is no redundancy in this model, as each element has an ID number attached to it, uniquely identifying it within the model. This model enforces certain rules, such as ligatures containing only notes, and a ligature having to include at least two notes. There is the possibility for data reuse, as the data in the graphical domain could also be used to answer questions about scribal practice, for example. This influences the stability and flexibility of the data model, but as the corpus of ND notation is closed (i.e. no new notation is being created) the flexibility of the model is not of great concern. The model is elegant as it requires only enough data to create a useful tool and models graphical elements as simple shapes. The model is also easily communicable as each musically distinct item is modelled separately, and any ambiguity is editorially worked out beforehand.

[5]For example, Payne (2021–) transcribes these duplex longas with a red line over the notehead to signify their length. At the opening of *In columbe* (O4) **F** f.67v (Payne 2021–, v.5, p.14), Payne applies these to some of the opening notes. However, only a few pages later he does not give the same treatment to the first notes of *-tum* in *Et respitientes* (O7) **F** f.70r (Payne 2021–, v.5, p.29), which are just as physically extended, if not longer. See also Apel (1941, p.245), who states that 'there is hardly any clear notational distinction' between simplex and duplex longas.

3.4.2. Analytical domain

The analytical domain is more complex to model as, unlike the graphical domain which attempts in part to capture the notation as it appears on the manuscript surface, the analytical domain is modelling virtual elements that are not related to some physical artefact but instead by conceptual links. Moreover, the shape of the analysis itself is dependent on the type of analysis being done. Suffice to say, an analysis of musical reuse in the ND repertory is going to be structured very differently to a harmonic analysis of the *Hammerklavier* sonata. As the previous chapter argued, it is not known with any certainty what the full picture of musical reuse in the ND repertory looks like, how large the set is, is or even what any of its elements are. However, we can assert that such a set exists, that each identifiable portion of musical reuse is made up of uniquely-identifiable passages of polyphony, and that the set of musical reuse can be found among the settings of music. If the same set of uniquely-identifiable polyphony is found in at least two different places, then that can be defined as a pattern of musical reuse.

More formally, this same plain-English idea can be expressed in terms of set theory, and the formalisation of this process will form the basis of the methodology outlined in “Chapter 6: Offline”. M is a set of uniquely-identifiable passages of music that are candidates to be considered as patterns of musical reuse: $M = \{m_1, m_2 \ldots m_n\}$. S is a set of settings of music that can be partially represented by all subsets of M, such that $S \not\subset \wp(M)$ but that there is also overlap between the settings' union set and the music power set:

$$\{S_1 \cup S_2 \cup \ldots \cup S_n\} \cap \wp(M) \neq \varnothing$$

We can then define the cartesian product of the settings, $T = S \times S$, and our set of musical reuse patterns as:

$$R = \{\ (\mathrm{f}(x) : x \in T) \,|\, (\mathrm{f}(x) \cap M \neq \varnothing\ )\ \}, \text{ where } \mathrm{f}(a, b) = a \cap b$$

What remains, then, is twofold: first a method to generate candidates (M), and secondly a way of representing settings of music (S) that could be recognised as subsets of M. A solution to this will be demonstrated later, however, it is currently sufficient to say simply that the representation of the analytical output domain can be represented as a set of sets.

3.4.3. Domain representational structure

Now that I have defined the domains of the input (graphical) and output (analytical) stages, I can define how and through what media the information is transformed. Defining this precisely will allow me to begin to ascertain what losses may occur along the way as well as, after McLuhan's theory of media, how the media will frame the

Representational structure of the domains of this study

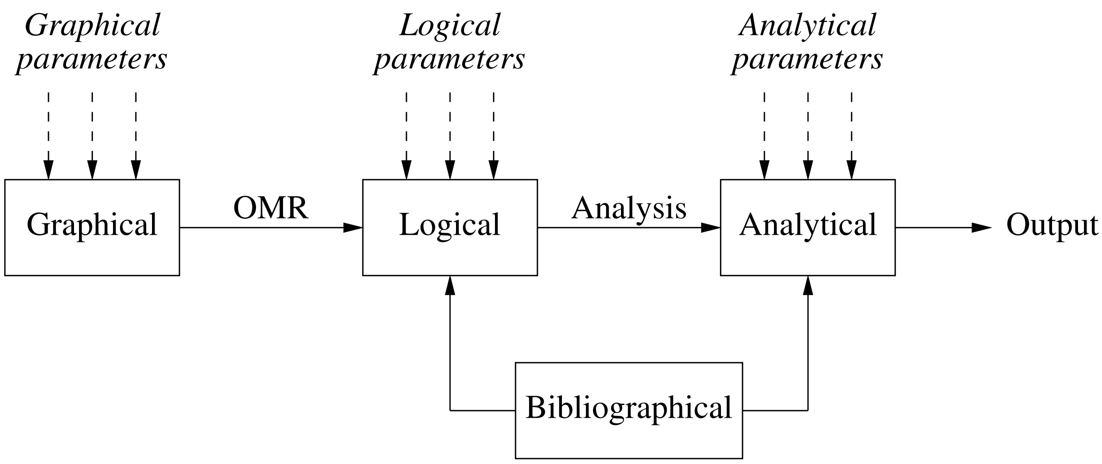

Fig. 3.6

message. "3.3. Representational structure" demonstrated that different encoding schemata attempt to relate domain through different routes. For example, Huron's (1992) conception would relate the graphical domain to the analytical domain through the nexus of CWMN, whereas Haus and Longari (2005) would link them through a logical domain. Baratè et al (2019) would take a more free-form approach, arguing that the two domains could be related directly or through any number of mediating domains.

This study will not adhere to one pre-existing representational structure but instead make the limited argument that in this particular case — that of linking the graphical to the analytical in this analysis for this repertory — the only possible course of action is to link through a logical domain (see Fig. 3.6). This is not the only way to link these two domains together — they could be linked any number of ways — but in this case a link from the graphical to the analytical through the logical is the only way to extract the necessary data from the notation. This is because the parameters already established in the graphical domain (such as *xy* coordinates) will not be good predictors of musical reuse, but parameters in the logical domain will. The sources in question were written over a significant period of time in different places by numerous different hands. Where one scribe may have written a passage of music in one line in a compressed hand, another may have broken the passage over multiple lines in a different clef. The parameters that define musical reuse are therefore not graphical but symbolic (e.g. pitch, ligation, and order) and so the graphical domain must first be interpreted into the logical to be analysed. As a result, it is not possible to take the data from the graphical domain and create an analysis directly: for the purposes of this study, musical reuse is encoded in the semantic meaning of the notation rather than its direct presentation in the manuscripts' notation. To complete the picture, it is then necessary to define the parameters of the logical domain for this analysis.

3.4.4. Logical domain

The parameters inserted at the logical domain can be defined as parameters that are necessary for the analysis (the output of this stage) that cannot be interpreted from the graphical domain (the input). Assuming that there exists a perfect system to convert the graphical data present within a stave of music into a symbolic representation (this could either be automated such as OMR or the result of human efforts),[6] the question that must be answered is: what is necessary for the analysis that cannot be interpreted from the graphical domain? Fig. 3.6 shows two sets of parameters other than the graphical domain feeding into the logical domain.

The first parameters are those from the small bibliographical domain already mentioned. This includes a knowledge of where settings of music begin and end, the names of these settings, how they relate to one another, and how to address them within the system (such as with an ID number). This data will be useful for querying the dataset and annotating the output with human-readable labels.

The second set of parameters are logical parameters introduced by a human editor. An OMR system could transform the graphical notation into pitches with the reasonable assumption that the music moves left to right. However, as described in "1.4.1. Layout and notation", the order of staves and systems does not always proceed from top to bottom of the writing block. One logical parameter is therefore a system for ordering staves and systems. Another logical parameter based on contextual relations is alignment between staves within systems. The surety that can be established in this alignment depends on the style of music. In syllabic *cum littera* notation and strict modal rhythm that moves in clear *ordines*, establishing a clear alignment between parts is rarely an issue. However, there are two main categories where alignment becomes difficult.

Firstly in modal rhythm, the exact rhythm intended from the grouping of ligatures is at times unclear, or a difficult hocket passage may misalign voices. In some cases, the strict patterning and/or the scribes' diligence in placing *divisiones* (as in Fig. 3.7), is enough context to infer the intended alignment, but this is not always the case. Secondly in organum purum, the alignment can only be indicated graphically, by vertical coincidence and *divisione* use, but more often than not this is unclear and must be inferred through an editorial process. For example, in Fig. 3.8, the flourished initial for the tenor *Benedicamus domino* has shifted the first note to the right. However, the organal voice begins above the initial, flush with the left margin. The first note of the

[6] With full knowledge after this chapter's discussion of McLuhan that such a "conversion" is impossible in reality.

W$_2$ f.1v. Hocket in modal rhythm

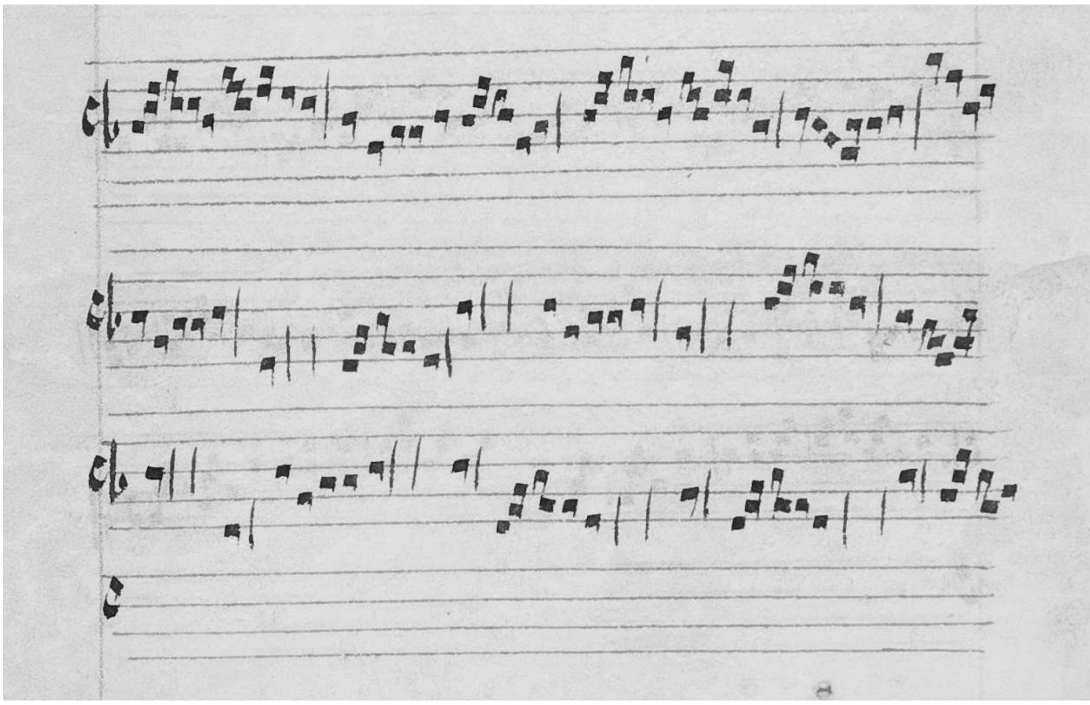

Fig. 3.7

F f.86v, I. Unclear alignment between tenor and organal voice

Fig. 3.8

tenor appears then to be coincident with the third *ordo*, not the first. Even more subtly, the third tenor note (a C) does not appear directly under the first C in the organal voice, aligning closer to the centre of the *ordo*. This alignment cannot be interpreted from the graphical domain but requires a logical parameter controlled by an editor that can read these contextual clues.

This same issue of alignment was dealt with directly by Erickson (1970), who created two basic principles for his study:

1. Changes of tenor note will normally occur at the beginnings and ends of phrases [*ordines*], rarely in the middle;
2. The last note of a phrase will form a perfect consonance with the tenor; a phrase may begin with a perfect consonance or ornamental figure that prepares a perfect consonance between duplum and tenor (Erickson 1970, p.50).

W_1 f.160v, I. Necessary editorial additions

Fig. 3.9

He then augmented this with a synopsis of the academic disagreement following Waite's (1954) controversial alignments between parts but came to no firm conclusions, summarising that these problems 'will require more detailed study of both organum and clausula than this study represents' before formulating a reasonable heuristic to determine an acceptable alignment (Erickson 1970, p.53). Karp (1966) has discussed this same issue in his own editorial practice for ND polyphony. Suffice to say, the issue remains unsolved.[7] Despite Erickson's efforts, his division of ND polyphony into neat "phrases" is far too simplistic: even in duplum, tenor and organal voice regularly overlap each other. Alignment cannot be inferred solely from the graphical domain and requires some editorial input.

Debussy, *Danseuses de Delphes*, bb.7–8 (1910)

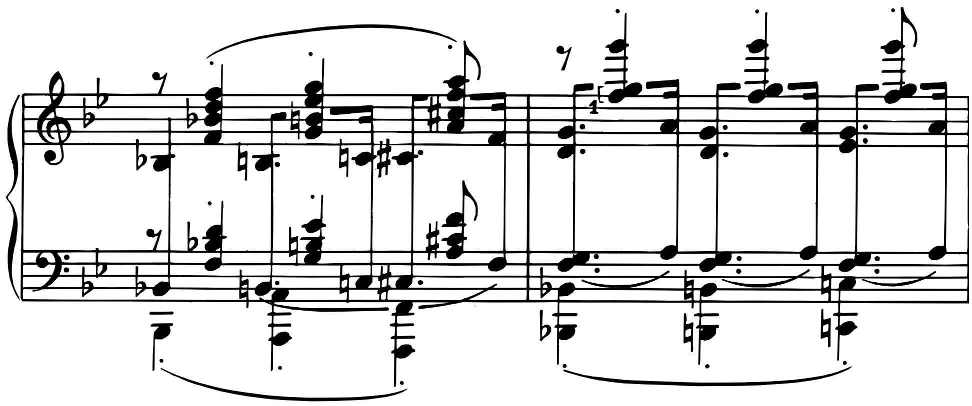

Fig. 3.10

[7] Erickson's study has been unfairly treated in the literature, both by historical unavailability (Karp 1992, p.51) and an inaccurate statement concerning his treatment of consonance: 'Erickson's classification (buried on p.77 of the computer printout) of the sixth as a consonance' (Yudkin 1983, p.359). Erickson's code does not classify the sixth as a consonance: it ranks thirds and sixths at a consonance level of "1", i.e. less consonant than fourths, fifths, octaves and unisons, but more consonant than seconds or sevenths.

3.5. Data model design decisions

This is without mentioning the issue of scribal errors and editorial additions. In Fig. 3.9, (a passage from the conductus *Deduc syon uberrimas*) the lower voice may look correct, but the flat added at the end of the first *ordo* could be interpreted many different ways. It could be interpreted as a G-flat (which would be incorrect) or a re-naturalisation of a previous F-sharp (there are none). It could be noticed then that the third note of the second *ordo* in the lower part (an F in this clef) clashes with the upper part (a G). It is clear here that the scribe has neglected to write a new clef for the second *ordo* and the flat is in fact a B-flat, and should read CBBGAGFE. However, in the final *ordo* of the line, continuing in this clef begins the *ordo* on a D against a C. Changing the lower voice back to its previous C3 clef does not help. It is only by looking at its concordances — such as **F** f.336v, III and $\mathbf{W}_2$ f.94v, II — can it be determined that the final *ordo* of the line in the lower voice should be written in a C4 clef, reading FGABC. Here, too, editorial input is required and the graphical domain itself is not sufficient.

Now that the data requirements for the analysis have been created, a high-level model as well as data structures for creating, storing and transforming this data can be formulated so that the creation of analyses is made as efficient as possible. However, it is important here to balance my own uses against Simsion and Witt's criterion of data reusability (2005, pp.11–12). I should design the model and data structures such that they could be repurposed for other analyses: at every stage the processes should be broad enough that they could be easily turned to ask other questions. This is another advantage of beginning from the "raw" graphical domain: the logical parameters have not yet been enforced on the data and another researcher could easily use a dataset from the graphical domain with little concern for bias. The graphical domain will be represented as a database storing the parameters listed in "3.4. A framing for ND polyphony", alongside user tools for input such as converting, querying, and interpreting the data into the logical domain. The analytical domain will take form as a set of patterns of musical reuse, each containing a yet unknown number of parameters. A computer program could then be made to "read" staves encoded in the graphical domain, combine this information with the logical parameters already set out — such as alignment and editorial markings — and create a logical data structure. The logical data could then be combined into a corpus of ND polyphony, and an analysis performed on the corpus to extract uniquely-identifiable passages, combining them in sets using the logic outlined in "3.4.2. Analytical domain".

The motivation behind the design decisions just outlined are relatively simple to understand: there is an element of necessity in their design due to the attendant necessities of the input and output stages. On the other hand, the logical data structure requires more careful selection. Firstly, it is the transformative layer between the input

W$_1$ f.32r, I

Fig. 3.11

From Karp (1966)

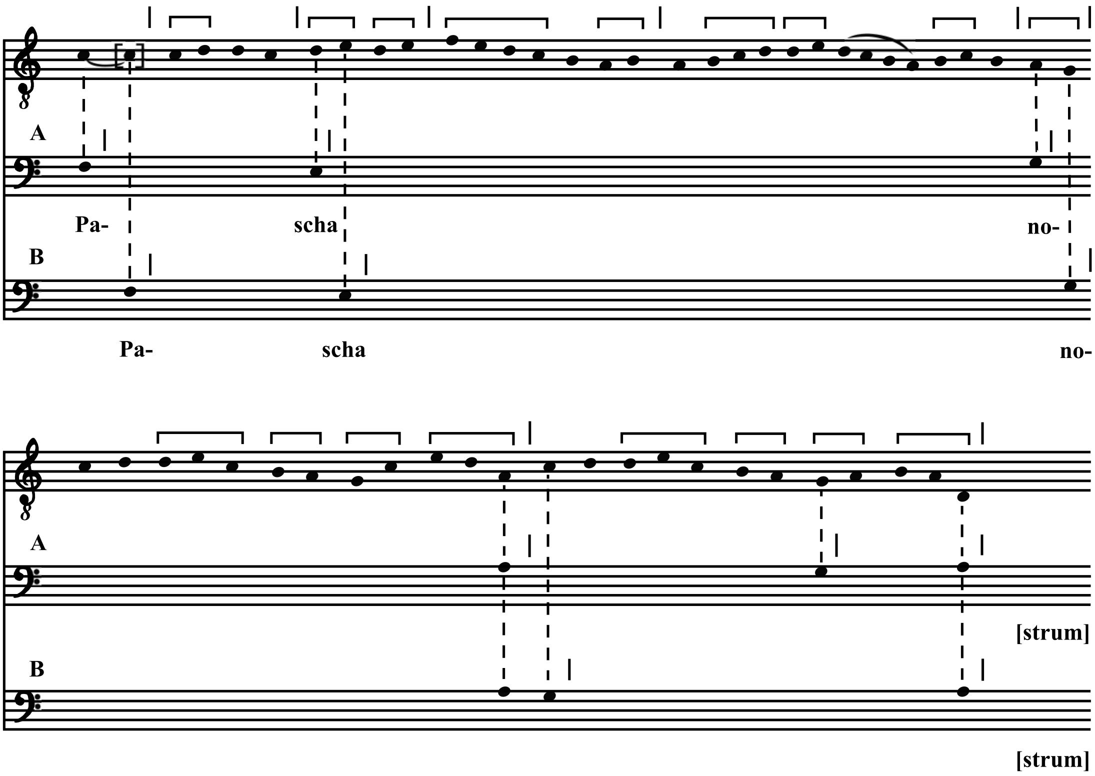

Fig. 3.12

Naïve vertical alignment of the opening of *Crucifixum in carne*, **F** f.71r, VI

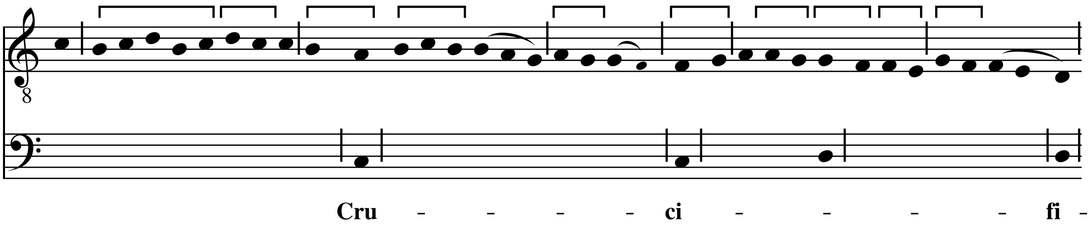

Fig. 3.13

graphical and output analytical domain. Secondly, it is a virtual domain, having no related artefact and it is therefore important that this domain is tightly connected back to the graphical domain so that its transformation is rooted in the notation and not some vague concept of how I might wish the music to operate. Finally, there are numerous overlapping and competing standards in the encoding of the logical domain which must be evaluated for their suitability in this case.

3.6. Encoding attempts

I have argued elsewhere that the three most useful data models for music encoding are tables, trees, and graphs, and that most if not all standards for music encoding fall into one of these three categories (Stutter 2024). Encodings either attempt to align vertical simultaneities together in tables, create hierarchical models (trees), or link items together in directed graphs. Byrd (1984) established the requirements of his computer music typesetting system by selecting difficult passages of music and ensuring that encoding them was possible in the system. For example, he selected a passage from Debussy's *Danseuses de Delphes* (*Préludes* 1910, no.1) as an example of how a data model of beams in CWMN cannot rely on the assumption that they are always unbroken (Byrd 1984, p.61): in this case, the cross-staff beams are interrupted by interjecting chords in the right hand (Fig. 3.10). This guaranteed that his final result was flexible, not only to accommodate the most common tasks, but also to cover some of the more esoteric notations that he wished to support. Although Byrd's study did not cover ND notation, only CWMN (Byrd 1984, p.1), his process for finding the limitations of certain models through numerous wide-ranging counterexamples is a sound method for discovering the exact domains and relationships that the model must support. Therefore, this section has similar aims, albeit on a smaller scale, i.e. solely within ND notation. I will therefore select three passages of music that others have noted as containing modelling difficulties, attempt hypothetical encodings as tables, trees, and graphs, in order to evaluate the success of each encoding for this study's purposes.

Firstly, Karp (1966, p.351) highlighted just such a tricky passage in organum purum from $\mathbf{W}_1$'s transmission of *Alleluya. Pascha nostrum* (Fig. 3.11) with alignments at odds with one another from Ludwig and Bukofzer respectively (Fig. 3.12). This is not surprising, given Karp's observation that an over-reliance on the vertical alignment as present in the manuscript sources 'would result in a harmonic structure that might be described as irrational or even idiotic' (Karp 1966, p.353). Karp did not provide any examples of this, perhaps because the disagreement between Ludwig and Bukofzer in *Pascha nostrum* is striking precisely due to $\mathbf{W_1}$'s relatively clear vertical alignment at this point. However, for a clearer illustration, we can briefly consider a naïve vertical alignment of the opening of *Crucifixum in carne*, from **F** f.71r, VI (Fig. 3.13). In this figure, the tenor is placed exactly where it aligns vertically in the manuscript, but it is clear to see that in many cases this is clearly not the harmonic structure that was intended by the scribe (most obviously that the tenor does not begin at the same time as the organal voice). Therefore, any conversion from the graphical to the logical domain requires some editorial parameters to achieve a satisfactory harmonic structure, based on striking a careful balance between structure and vertical alignment. Both Ludwig and Bukofzer's editorial alignments arguably do contain that balance but, over half a century later, exactly where that balance lies is still unknown.

Secondly, even where there is some semblance of modal rhythm, in *modus non rectus*, different transmissions of the "same" polyphony contain different ligature groupings. Roesner (1981) supplemented an example taken from Reckow (1973) (Figs. 3.14, 3.15) to demonstrate the re-ligation — and therefore to his mind the likely rhythmic reinterpretation — of the same music in the three central sources, something that he attributed to 'the house style or a particular institution or cantor' (Roesner 1981, p.375). The encoding of ND notation for this study should therefore not rely on a single rhythmic interpretation as this may reflect a reliance on a single manuscript as a base text, privileging a single interpretation. Moreover, the rhythmic interpretation of ND notation is often extremely subjective — a fact first raised by Parrish (1957, pp.87–91) — and therefore the approach used here should not rely on rhythm as a requirement for encoding. This is because the rhythm implied by the ligature groupings may not be known, or the exact interpretation contentious. The three versions can be seen in manuscript in Figs. 3.18, 3.19, and 3.20.

Finally, there are passages in strict modal rhythm where the non-standard configuration of ligatures makes the rhythmic interpretation and therefore synchronicity of the voices ambiguous. Baltzer (1974, p.98) described *Tamquam* No. 9 in **F** (Fig. Fig. 3.17) as 'full of irregularities and ambiguities, with frequent fracturing', and it being 'often difficult to tell whether the scribe is writing a 3 li [ligature] or a simplex and a 2 li (or a 4 li as opposed to a 1 + 3 li)' (Baltzer 1974, p.99). Baltzer provided multiple interpretations of the same passages where the exact alignment of the notes is altered depending on the rhythmic interpretation of the reader (see Fig. 3.16).

From Roesner (1981)

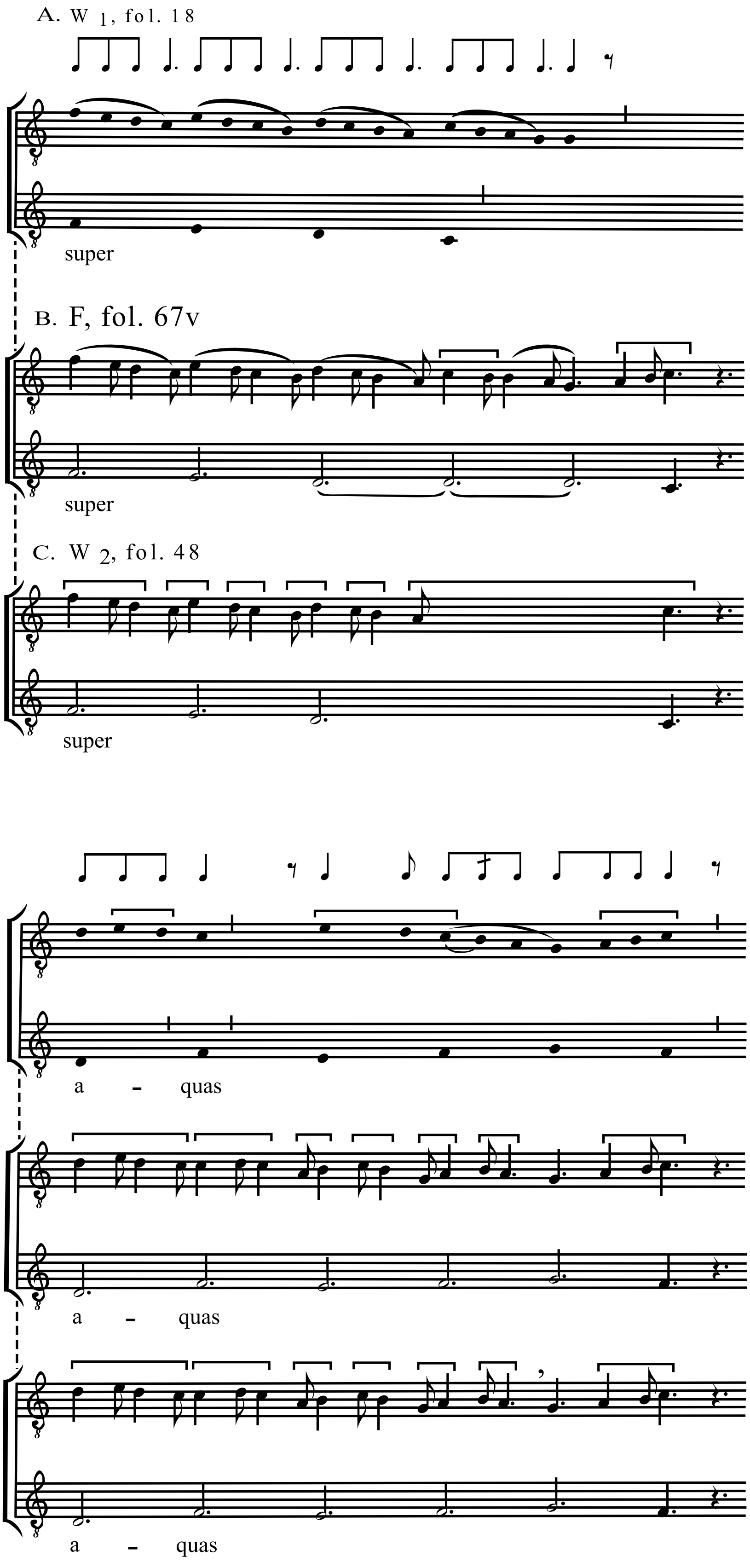

Fig. 3.14

From Roesner (1981) (cont.)

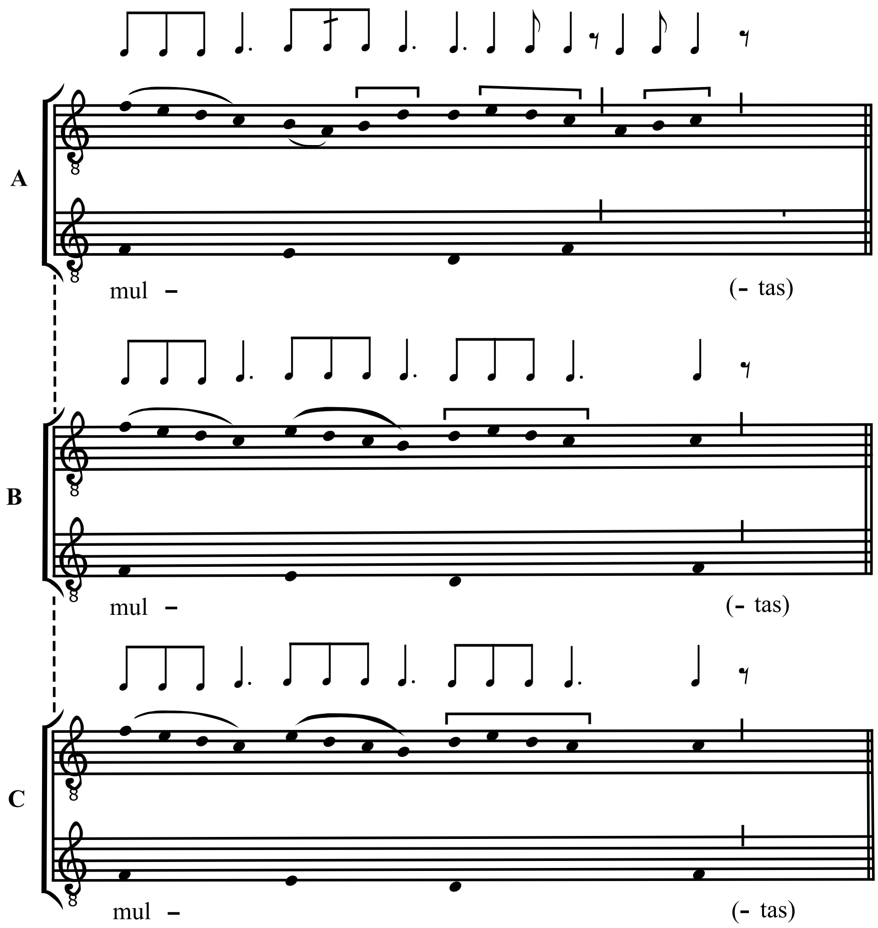

Fig. 3.15

Tamquam no.9, adapted from Baltzer (1974), with alignment between concordances

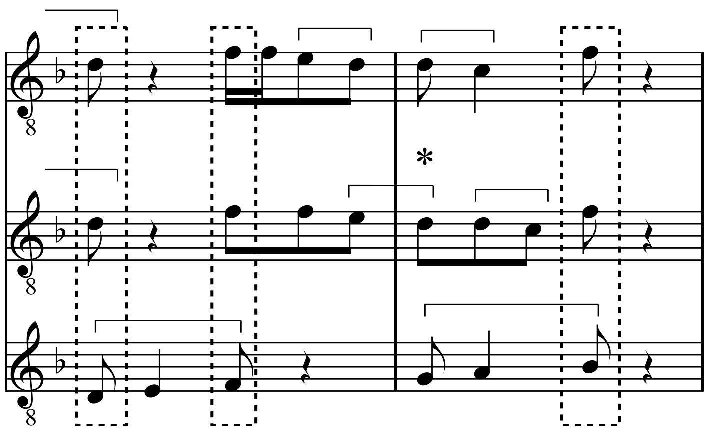

Fig. 3.16

F f.148r, II

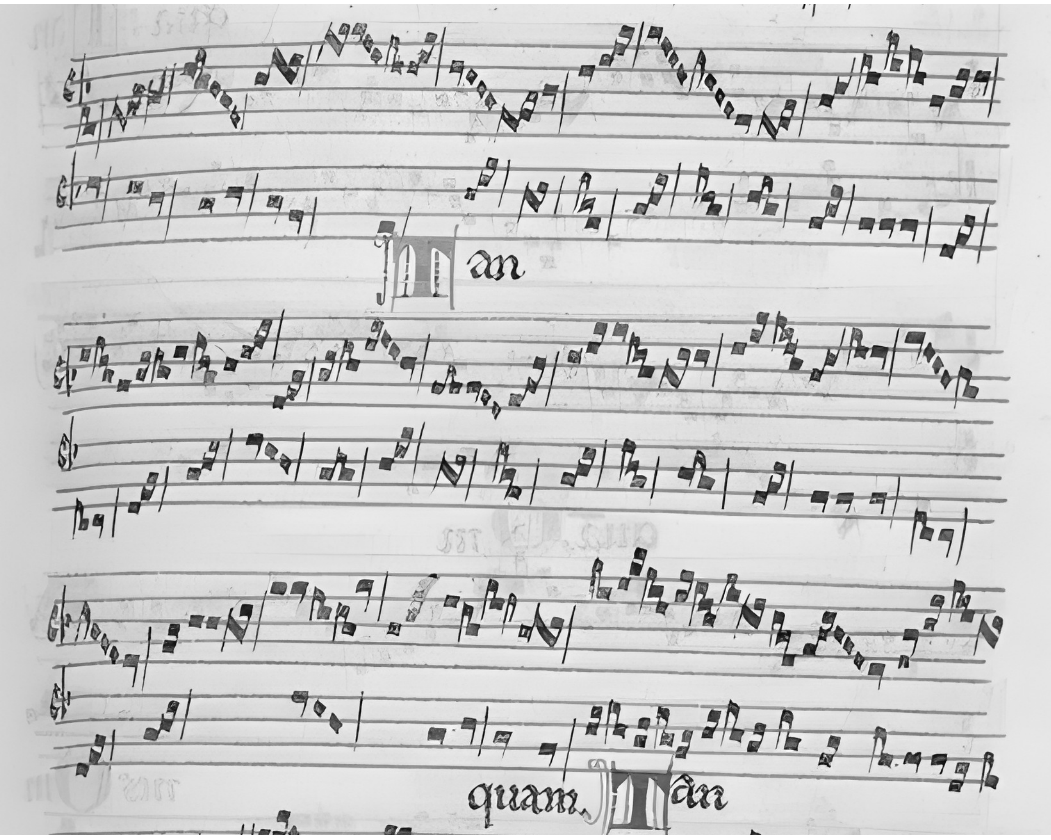

Fig. 3.17

W$_1$ f.18r, VI

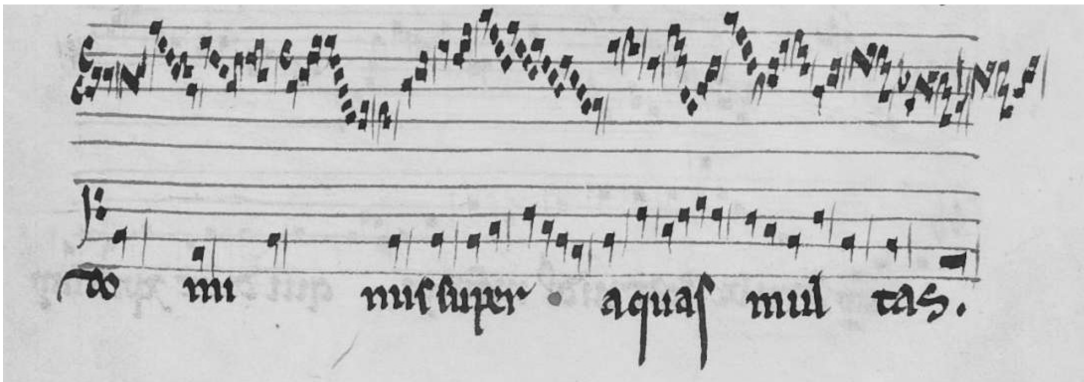

Fig. 3.18

Therefore, the encoding in this study should be able to reflect the possibility of notes being aligned with more than one other note.

I therefore have three tricky test passages to each be encoded as tables, trees, and graphs: the opening of **W$_1$**'s *Pascha nostrum*, three parallel transmissions of *Super aquas mul[tas]* from *In columbe*, and *Tamquam* No. 9 from **F**. It is important to note here that very few specialist schemata for ND polyphony exist, so hypothetical extensions to already-existing schemata may be necessary.

F f.67v, IV–V

Fig. 3.19

W_2 f.48r, IV

Fig. 3.20

3.6.1. Tables

The tabular representation format that I will be using for demonstration here will be *Humdrum*, which is described in Huron (1995, *Humdrum* does not have a fixed set of representation systems, although it does provide a number of useful defaults, such as the `**kern` representation scheme for CWMN.[8] Extending *Humdrum* is informal and encouraged (Huron 1998, p.2), and so creating an ad hoc representation scheme for ND polyphony is appropriate here. Other medieval repertories that would otherwise not be able to represented using CWMN notation have been represented using specially-designed *Humdrum* syntax, for example Morent (2000) designed the `**Hildegard` representation for the particular scribal practices of the Hildegard manuscripts. It is therefore possible to design a *Humdrum* representation, based on the approach taken by Morent (2000) of separating pitch information from ligature information, and combine

[8] Representations in *Humdrum* are known by their "spine" (column) names, which always begin with two asterisks.

it into a polyphonic *Humdrum* score (Figs. 3.28 and 3.29).

This encoding imagines a hypothetical `**NotreDame` representation complementing the more typical `**kern`, where various characters indicate particular ND polyphony events: square brackets for square ligatures, parentheses for *currentes*, a pipe character (`|`) for *divisiones*, and a caret and lowercase letter V (`^`, `v`) for upward- and downward-facing *plicae*, respectively. The parallel spines represent the organal part as well as the alignment given by Ludwig and Bukofzer as indicated in Karp (1966). Such a representation is successful, although extremely verbose. This can then be expanded for (optional) rhythmical interpretation by appending `L` or `B` to the annotation, indicating a longa or breve respectively. Perfect longae are be indicated by adding a dot (i.e. `L.`) and longer notes by adding a multiplier (such as duplex longae by `L.2`).[9] The figures from Roesner (1981) can be represented using this schema (Figs. 3.30 and 3.31).

The success of the tabular data representation unfortunately becomes untenable when we begin to consider the alignment issues in the example from Baltzer (1974). The competing rhythmic interpretations of the ligature patterns cause issues with alignment. At the asterisk (*) in Fig. 3.16, the upper rhythmic interpretation places the first D as coincident with the last breve of the previous perfection, and coincident with the rest in the tenor indicated by the *divisione*. However, the lower interpretation places that same note with the first breve of the next, coincident with the G in the tenor. Both interpretations are equally valid: there is not a case of a “right” or “wrong” answer in this case, and this is part of the reason why Baltzer here explicitly stated both possibilities. Turning to our tabular representation, the rows of the table are a sequence of exact coincidences: every row is a discrete and exact moment in musical time. This cannot be reconciled with the possibility that the rhythmic interpretation may be ambiguous, and so examples such as these cannot be represented tabularly. All that can really be agreed upon is that some moments are aligned (see the dashed boxes in Fig. 3.16). However, there is no way in a tabular representation to express the common denominator between both interpretations: that between two perfections there are a series of notes, but being unable to define exactly how they actually line up.

3.6.2. Trees

A hierarchical data model for encoding ND polyphony exists. As part of my Master’s thesis (Stutter 2020), I designed a format called *ndp* that solves the very issue the tabular format has,[10] for example the issue with the *Tamquam* passage could very well

[9] The durations cannot be encoded directly in `**kern` and must be represented in the `**NotreDame` representation as, to the best of my knowledge, `**kern` does not have a rhythmic specifier for longae (either simplex or duplex).

[10] The entirety of the organa of $\mathbf{W}_1$ are encoded in this format at <https://gitlab.com/yockyrr/w1-ndp>.

be encoded as:

```
{{ff[ed][dc]}{f]|[ga}}
{{f|}{b_]}}
```

Each discernible group (or *ordo*) of music is encoded within consecutive sets of braces, and ligatures are indicated using square brackets. This encoding therefore indicates that the note sequence FFEDDC fits within a space where the tenor has FGA. An unsupported oversight within *ndp*, however, is overlapping notation, such as in hocket, as the groupings (*ordines*) created by the braces (`{` and `}`) must be self contained.

We could instead imagine a better version of *ndp*, one that was built as an extension to another common and well-supported hierarchical format such as the *Music Encoding Initiative* (MEI). MEI is an XML-based encoding standard that places musicological concerns as its central remit. Unlike other standards such as *MusicXML* which place interoperability between notation software as their primary concern (Crawford and Lewis 2016, p.282), MEI has first-class support both for various domains of encoding and for extensibility, whereby the standard can be formally customised to change, add, or remove elements. Such features are indispensable for non-CWMN notations as found in early music (Roland et al 2014, p.609), and the XML format allows for both human- and machine-readable encodings. An ND polyphony extension could very well be constructed that, like the *Humdrum* experiment, added support for ND polyphony elements such as ligatures. On the surface, this sounds like a fine improvement: however, the structural issues of overlapping notation and alignment would still remain.

To demonstrate these structural issues, consider a CWMN fragment in MEI representing a simplified version of the opening of *Sumer is icumen in*:

```
<section>
  <staff n="1">
    <layer>
      <note oct="5" pname="c" dur="4"/>
      <note oct="5" pname="c" dur="8"/>
      <note oct="4" pname="a" dur="4"/>
      <note oct="4" pname="b" dur="8"/>
      <note oct="5" pname="c" dur="4"/>
      <note oct="5" pname="c" dur="8"/>
      <note oct="4" pname="b" dur="8"/>
      <note oct="4" pname="a" dur="8"/>
      <note oct="4" pname="g" dur="8"/>
    </layer>
  </staff>
  <staff n="2">
    <layer>
      <note oct="4" pname="c" dur="4" dots="1"/>
```

```
        <note oct="4" pname="d" dur="4" dots="1"/>
        <note oct="4" pname="c" dur="2" dots="1"/>
      </layer>
    </staff>
  </section>
```

How does a parser know that the D in the second stave is coincident with the A in the first stave? There is nothing in the structure itself that links them together. Rather, for each stave the parser must keep track of time, and place coincident notes together.

This is no problem for music notation with fully-specified durations or monophonic music where alignment is not an issue. However in ND polyphony, durations are not fully specified but alignment is key. The hierarchy could be flipped to be more conducive to vertical coincidences, for example, by creating a series of time instance elements and placing each stave's notes inside. This second possibility is only a minor improvement on the tabular structure and is structurally identical to *ndp* which, as already explained, has the opposite problem of an unsatisfactory representation of overlapping notation. This problem of single-hierarchy structures causing problems for multiple views on music, especially overlapping phrases has been known since Dannenberg (1986), and the application of this critique to XML structures such as MEI can be seen at its fullest in Wiggins (2009a, pp.20–21).

3.6.3. Graphs

Finally, then, there is the option of a graph data model. A graph data model is inherently more flexible than tables or trees as its links are undifferentiated in syntax: a complex or esoteric link (such as complex alignment) is made by the same mechanism as a simple or common link (such as one note following another). This comes at the expense of human readability (Stutter 2024). Graphs are also intensive to parse: as the next link in the structure could be anywhere else in the entire graph, it cannot be parsed as a stream (such as *Humdrum*) or as a series of events (such as MEI using a SAX parser). Each element of the graph data structure must therefore be held in memory which may become a problem for extremely large graphs. With so many drawbacks for usability, a graph data structure often becomes the data structure of last resort when, as in this case, the options of tables and trees have been seen to be insufficient for this study's purposes.

Despite being a "tree" XML structure, MEI does also contain some graph-like features. Often, the schema affords links to other elements by specifying XML IDs or even URIs. This is inherited from MEI's immediate ancestor, the *Text Encoding Initiative* (TEI), where links are used to 'represent analyses of the structure of a text which are not necessarily linear or hierarchic' (Text Encoding Initiative 2025).[11] In MEI, the attributes `@copyof`, `@sameas`, `@corresp`, `@next`, `@prev`, `@precedes`,

@follows, @synch, and @when can be used on many different elements to provide these kinds of links.

In this case, the @synch attribute can be used to encode the varied alignments between Ludwig and Bukofzer, with MEI's <app> and <rdg> critical apparatus to encode the differences themselves. This can be seen in Figs. 3.32– 3.35, within a proposal of how ND polyphony may be encoded in MEI. Similarly, Roesner's (1981) figure can be encoded by co-opting the @num and @numbase attributes from the MEI mensural module to encode implied rhythm, and @corresp to link notes between readings. This can be seen in Figs. 3.36– 3.45. Baltzer's (1974) figure could be encoded any number of different ways, either by using alternate readings using MEI's critical apparatus or dispensing with rhythmic interpretation entirely and encoding each voice separately with @synch links, as I have done with Roesner's figure.

3.6.4. Encoding conclusion

The conclusion from this short study is that to efficiently encode the logical domain of ND polyphony, a graph data model is required, as both tabular and hierarchical data structures misrepresent the notation, particularly in overlapping notation and passages of ambiguous alignment. A graph data structure, perhaps using the graph-like features of certain MEI attributes, will be the best format to represent the logical domain of the notation.

With respect to MEI there are some downsides, however. The first is the extreme prolix of the syntax: the tabular representation, although insufficient, was much more terse and readable. This makes the syntax of XML-based markup such as MEI difficult to write, as each element may be linked to and therefore requires its own @xml:id. These IDs must be generated and carefully copied to create the graph structure. This is very difficult to do manually and therefore, where possible, the XML should be generated automatically. Secondly, although MEI defines its own ontology in the linking attributes already described, this is the extent of the linking capabilities without also extending the att.linking attribute class to be more generic. For example, it is currently impossible to apply the critical apparatus tools (<app>, <rdg> etc.) to the linking attributes, and it is unclear how this may be achieved elegantly.

MEI's critical apparatus tools are wrapped around entire elements, and so if the difference between two readings only concerns the value of one attribute, the entire containing element must also be replaced. Not only is this inefficient but, for those attributes that are not changed, replacing the attribute with an identical attribute is not the same as not replacing it: for example the elements cannot have the same @xml:id

[11]This phrase has been part of the TEI guidelines since 2005. See <https://github.com/TEIC/TEI/blob/8ed19c0a616e4a816e3255629209d1e9687a1c51/P5/Source/SA/sa.odd#L29>.

and must then be addressed individually. Moreover, the hierarchical structures of MEI's critical apparatus would not gel well with the intended graph structure. What would be required to solve this issue would be critical apparatus able to add or remove links in the graph: what this apparatus would look like, where it would go in the structure, and how it would operate is not known. Above all, these issues emerge due to the fact that, although XML is very flexible and extensible, affording all manner of data structures, it is designed primarily for hierarchical data,[12] and it is therefore to be expected that designing a system using XML for graph structures would be much more cumbersome than a native graph data format.

3.7. Polyphonic problems

With a symbolic graph representation for ND polyphony, the next step would be to then transform a whole corpus of these graphs into an analysis of musical reuse. How can such graphs be analysed to measure their similar features? For language, such techniques are contained under the heading of corpus linguistics, and the atoms of language such as words are analysed to generate concordances and collocations. The granular analytic process to be used here is broadly the same, but within the musicological constraints of ND notation. Although there has been significant research into applying the techniques of corpus linguistics to monophonic melody, there is no such methodological structure for polyphony. The reason for music's slower progress in this area has been well summed up by Quinn (2014, p.295), who claims that 'the excuse is technical: it is easy to transform literary texts into searchable data, and hard to do so for music'. In terms of data, texts are a one-dimensional structure: it is possible to split a text into a linear series of tokens (such as words) and, for example, vectorise the result for numerical analysis.

Similarly, monophonic melodies can be encoded as text with the same dimensional properties and tokenised in the same way. Repertories such as folk song and plainchant have been studied digitally this way since the mid-twentieth century (Bronson 1959). However, it is difficult to analyse polyphonic music in this same way as it is more difficult to tokenise the interplay of multiple voices into a contiguous stream of events. The current state of the art methodologies generally begin by "chordifying" the music, i.e. collapsing the multi-dimensional structure of polyphony down into a one-dimensional array of time slices by alignment (Sears and Widmer 2020, p.213). Given the previous section's findings that ND polyphony cannot be fairly sliced this way, due to ambiguities in rhythmic alignment between voices causing elements to not belong to one slice or another, such a methodology will not be appropriate for this study.

[12] XML is a subset of *Standard Generalized Markup Language* (SGML), which in its ISO specification makes repeated reference to hierarchical structure (ISO 1986).

In the previous sections that appraised data structures, all of the structures surveyed were capable of reasonably encoding the simplest examples, but some failed to adequately represent the more complex instances chosen. Nonetheless, it was possible to encode the necessary data satsifactorily as a graph model. How to encode ND notation in the computer therefore already has a conceptual answer, and all that remains is to create the logical data model using pre-existing tooling. However in the case of analysis, there is a distinct methodological infancy for approaching the novel research questions asked in this study, and it is unlikely that there will be an off-the-shelf solution. The possibility of detecting musical reuse is not a question commonly investigated within the study of music. A related inquiry, similarity search, has long been a core task in music information retrieval (Selfridge-Field 1998, p.4), but is typically viewed through an approach of clustering musical pieces into a known number of groups (Müllensiefen and Frieler 2022). Such methodologies are suitable for whole-piece questions of similarity, such as stylometric analysis (Cuenca-Rodríguez and McKay 2021).

In the case of ND polyphony however, we do not know how many groups there are, and these groups are likely to be small, many including just one or two passages. One-to-one similarity of shared music material is much rarer in CWMN, and so there has been little need to develop methodologies for finding such links. Projects that do consider direct similarities, such as the CRIM (*Citations: the Renaissance Imitation Mass*) project, typically consider similarity as a feature that is to be manually extracted by musicologists, with or without the help of computational tools, rather than entirely automatically (Morgan et al 2022). As was seen in chapter 2, manual similarity extraction in the ND repertory is variable and highly dependent on idiomatic criteria.

When added to the data structure complexities already discussed, this leads inevitably to the conclusion that analysing ND polyphony for musical reuse will require a ground-up investigation of possible methodologies, and even completely new methodological approaches. Therefore, before applying any analyses to the more complex problems outlined in the previous sections, I will first examine the methodologies available for the dimensionally simpler case of monophony. I will investigate whether these methodologies are suitable for this kind of analysis, and then whether they could be adapted for polyphonic notation graphs.

To investigate these methodologies, I need to identify some appropriate examples. To eliminate any confusion against the peculiarities of the ND repertory, key, or phrase length, I will appraise these methodologies using trivial examples of non-ND notation first, and later implement the most successful principles into ND notation and polyphony. I have therefore selected two monophonic examples from some well-known CWMN: the main theme from the final movement of Brahms' *Symphony No. 1* (see

Theme from Brahms Symphony No. 1

Fig. 3.21

“Alphorn” theme from Brahms Symphony No. 1

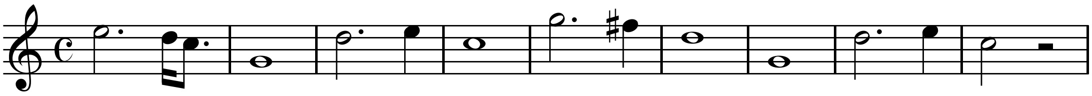

Fig. 3.22

Brahms theme annotated with numeric vocabulary

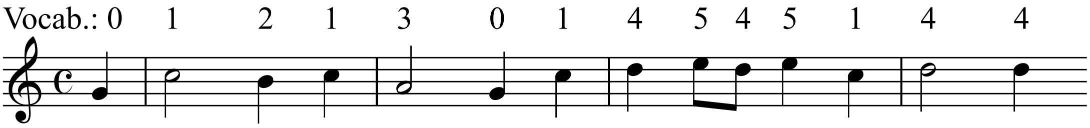

Fig. 3.23

Fig. 3.21). as well as the “Alphorn” theme from the same movement (Fig. 3.22). Both of these examples are in Common time, C major and fourteen notes long, and so they can be placed into new methodologies without concern for phrase length or key. At this small scale, we want to find a methodology that is capable of detecting similar passages within these two examples. Overall similarity is not useful, as the similarities may be extremely localised. These similarities may also be transformative, containing at one end of the spectrum of possibilities, only a general sense or gist of each other, and at the other, exact matches. All of these requirements must be fulfilled if the methodology is to be scaled effectively to the corpus of ND polyphony.

To my knowledge, there are four suitable methodological avenues to investigate:

1. **N-grams:** which enable the collection of local fragments of music together that can then be compared for similarity.
2. **Edit distance:** a set of related methodologies for quantifying the difference between strings of text.
3. **Sequence alignment:** another set of methodologies — this time taken from bioinformatics — for numerically calculating how two sequences could best be compared.
4. **Embedding and vectorisation:** a recent methodology heavily used in machine learning for calculating where elements exist semantically within a latent space so that they then can then be clustered.

3.7.1. N-grams

A typical technique from corpus linguistics is to create collocations based on n-grams. An n-gram methodology groups *n* tokens together in a continuously overlapping sequence that can then be used to analyse corpora structures larger than the probability of the token itself. The term was coined by Shannon (1948), who used it as a general term for the hyponyms “digram” and “trigram”. To give a simple example, Weisser (2016, p.196) introduced n-grams by the 2-grams “strong coffee” and “black coffee”. The adjectives “strong” and “black” are not meaningful by themselves but here tell us something significant about the coffee. The collocations “strong coffee” and “black coffee” are therefore significant where the individual words “strong”, “black”, and “coffee” are not. N-grams can also be used as a Markov model in text prediction, where the previous *n* tokens can be used to calculate the likelihood of the next token. There are innumerable successful uses of n-gram methodologies in corpus linguistics, and the methodology has become as common as more basic tools such as simple concordance and collocation, such that n-grams are introduced as ‘one of the most basic NLP [natural language processing] tasks […] elementary for any kind of statistical analysis’ (van Gompel and van den Bosch 2016, p.1).

More recently, n-gram methodologies have been generalised into so-called “skip-grams”, where an n-gram could be described not just by adjacent tokens, but by all tokens within a certain distance, effectively “skipping” some tokens. These were first described by Huang et al (1993) as “long distance” grams, and the co-ordination of a group of methodologies into the term “skip” grams was first introduced by Goodman (2001). Skip-grams have two advantages: the first is that the number of skip-grams that can be created far outstrips the number of n-grams due to the flexibility of the skip, and this is particularly useful in small corpora where the number of n-grams created may not be significant. The second is that larger syntactic structures can be determined from skip-grams without resorting to a larger *n* that would lose some of the more fine-grained detail.

For example, in the Harvard sentence “Glue the sheet to the dark blue background” (IEEE 1969, p.239), the bigrams would be “glue the”, “the sheet”, “sheet to”, “to the”, “the dark”, “dark blue”, “blue background”. However, 1 skip 2-grams would include all the tokens within one skip from the target token: “glue the”, “glue sheet”, “the sheet”, “the to”, “sheet to”, “sheet the”, “to the”, “to dark”, “the dark”, “the blue”, “dark blue”, “dark background“, “blue background”. We have almost doubled the number of grams extracted and can also now relate the adjective “dark” to the background and verb “glue” to the sheet without resorting to 3-grams.

Applying this to the monophonic instance of the main Brahms theme, I can first encode the pitches of the tune as a vocabulary of six notes in order of reading:

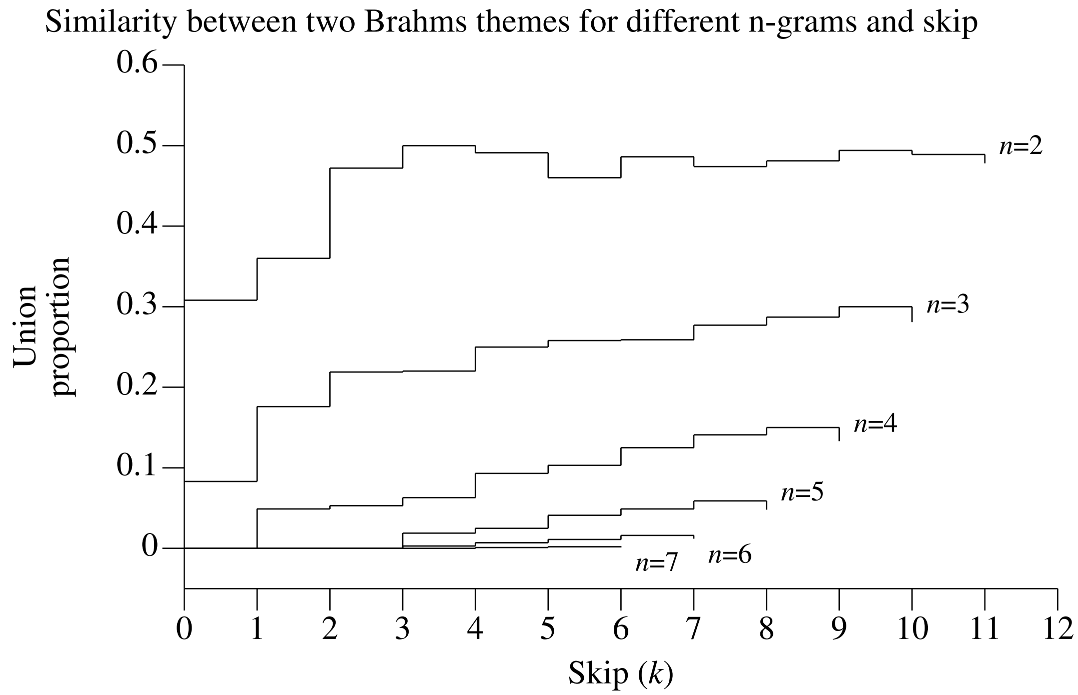

Fig. 3.24

01213014545144 (Fig. 3.23).[13] The list of 1-grams is then: (0, 1, 2, 1, 3, 0, 1, 4, 5, 4, 5, 1, 4, 4). We can see from this that all instances of 0 (the G) occur in the first half of the theme and all instances of 4 (the D) in the second. We can also consider 2-grams, yielding: (01, 12, 21, 13, 30, 01, 14, 45, 54, 45, 51, 14, 44) which can tell us something about the leaps between notes, and trigrams (012, 121, 213, 130, 301, 014, 145, 454, 545, 451, 514, 144) more about common melodic patterns.

Extending the methodology to skip-grams gives even more data, such as 1-skip-2-grams: (01, 02, 12, 11, 21, 23, 13, 10, 30, 31, 01, 04, 14, 15, 45, 44, 54, 55, 45, 41, 51, 54, 14, 14, 44). This begins to find relationships between items that are not adjacent to each other, such as the interruption caused by the neighbour-note quaver in the third bar. It is easy to see how extending this both in terms of corpus size (more melodies) and context (larger skips) can demonstrate more relationships, but increasing the skip also makes each result less significant as it becomes more likely that elements are near each other simply by coincidence. To compare the main theme with the "Alphorn" theme, I can extend the vocabulary with two more notes to encode it thus: 54104516740451. I can then see how many n-grams and skip-grams are in both the themes at different levels of *n* and skip (*k*) (see Fig. 3.24). For example, at 2-gram and 0

[13] These pitches could be encoded any number of ways, for example by scale degree (5, 1, 7, 1, 6, 5, 1, 2, 3, 2, 3, 1, 2, 2) or semitones from the tonic (-5, 0, -1, 0, -3, -5, 0, 2, 4, 2, 4, 0, 2, 2). However, in this case and to aid the transfer of the methodology to ND notation (a non-diatonic context) later, I encoded them here by assigning them ID numbers in order of reading precisely to avoid the misconception that the numbers given have any semantic meaning.

skip, the bigram “45” appears twice in both the first and second theme. However, at 2-gram 2-skip, the first theme extracts three instances of “45”, as the two instances of “45” in the first theme are next to one another and the increased skip allows the first 4 to also capture the second 5.

Depending on the parameters of n and k chosen, the similarity between the two themes using this methodology varies between its lowest value of 0% (e.g. n=5, k=2) and its highest value of 50% (at n=2, k=3). The main conclusion I wish to draw from this is that what we may define as “similar” depends greatly on the perspective, metric used, and granularity of the search. At a high n, these two themes can be considered completely dissimilar, but at a smaller n and correctly-chosen k, they share one in every two collocations. This can be seen as a digital corollary to the conclusions of chapter 2, where the criteria that are selected significantly alter the results. The difference here is that, whereas the criteria created in previous work are subjective and take years of study to apply to the entire repertory, the parameters of methodologies such as n-grams are exact and reproducible, and can be altered if necessary with little effort.

3.7.2. Edit distance

N-gram methodologies can be used to create statistical models of similarity within corpora, but to gain a more exact heuristic of the similarity between two specific sequences, we can use edit distances to score the changes required to transform one sequence into another. The most commonly used metric is the Damerau–Levenshtein distance, created from a synthesis of two scores (Hall and Dowling 1980). Levenshtein’s original paper created a distance score as the cumulative count of all insertions, deletions, or substitutions necessary to transform one sequence to another (Levenshtein 1965). The Damerau–Levenshtein distance also includes transposition of adjacent characters extended from the single edits Damerau considered into full strings of text (Damerau 1964).

For single examples, edit distances are much more effective than n-grams as they do not require the creation of an entire corpus. For the Harvard sentences “Glue the sheet to the dark blue background”, “It’s easy to tell the depth of a well” and “Rice is often served in round bowls” we can use the Damerau–Levenshtein distance to assess the sentences’ similarities between each other (Fig. 3.25) (IEEE 1969, p.239). Using this metric it could then be argued that the second and third sentences are more similar to each other than to the first sentence, due to their lower edit distance scores. In the case of the Brahms themes it is as simple as computing the distance between the strings “01213014545144” and “54104516740451”, which is 11 (3 insertions, 3 deletions, 5 replacements). Dividing this by the sequence length (14) gives us a per-token score: 0.79. Comparing the pairwise distance between different themes would be able to tell us

Damerau–Levenshtein distance between three Harvard sentences

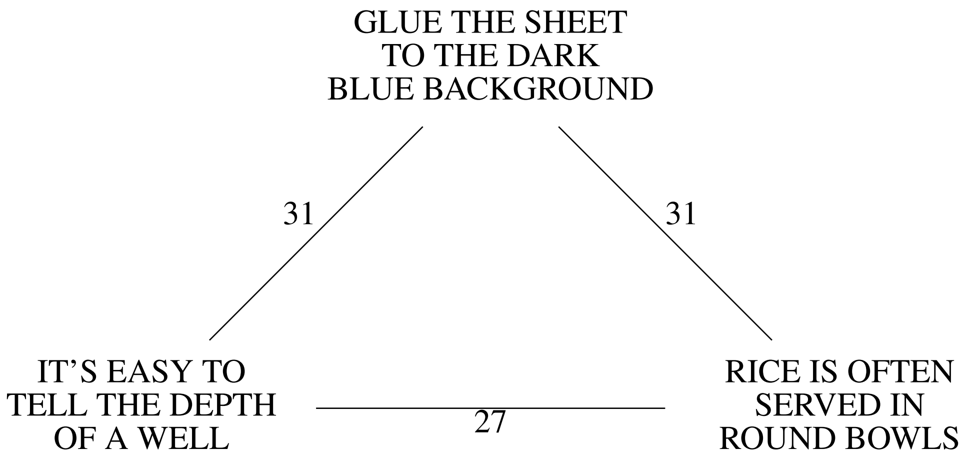

Fig. 3.25

Two Brahms themes aligned using Needleman–Wunsch

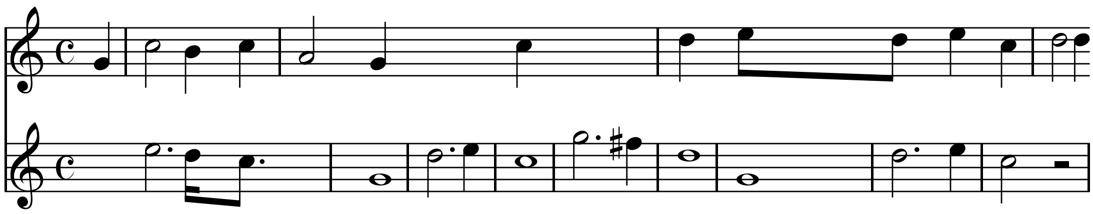

Fig. 3.26

more about whether this is a significant figure.

3.7.3. Sequence alignment

A related problem in bioinformatics is that of sequence alignment, e.g. finding regions of similarity between DNA and protein sequences. One well known algorithm for solving this is Needleman–Wunsch that determines, using a scoring metric, how best to align two sequences including "indels" — being places where the alignment does not exactly match — and therefore the most likely filiation between them (Needleman and Wunsch 1970). This is now a core methodology in bioinformatics and the intervening half century has yielded numerous improvements in the general technique. There exist multiple software packages that implement multiple sequence alignment for bioinformatics such as *T-Coffee* (Notredame et al 2000), *Seaview* (Gouy et al 2021), *MAFFT* (Katoh and Standley 2013), and *Clustal Omega* (Sievers and Higgins 2017). Multiple sequence alignment allows for the alignment of more than two streams.

Performing Needleman–Wunsch alignment on our two theme sequences yields:

```
012130--1--4545144
-541-04516740451--
```

There are seven moments where the alignment matches. This could be displayed in notation as in Fig. 3.26. Clearly in this case the alignment struggles to find any similarity between the unrelated themes, but it does demonstrate the soundness of the methodology for a related issue: that of aligning two musical sequences that are already known to be essentially the same with some minor variation.

3.7.4. Embedding

A final technique developed more recently is that of creating word embeddings: creating lower-dimensional vector representations of large vocabularies. This is a vectorisation method that can project a token into a d-dimensional space that can then be used to measure the similarity between tokens. Mikolov et al (2013) created the basic premise for modern language models by formulating the *Word2Vec* model that predicts the classification of a word either by the average of its vectorised context (Continuous bag of words: CBOW) or predicted surrounding words based on the vectorisation of that word (Continuous skip-gram). The following *Python* code demonstrates Word2Vec on the two Brahms themes using CBOW, then the projection of the vectors into two dimensions (from the original 5-dimensional space) using Principal Component Analysis (PCA) which can be seen in Fig. 3.27. The cosine similarity between the vectors could be calculated to quantify the distances between the items, or a clustering algorithm could place similar items together. For example, we can see from the plot quite clearly that '7' (i.e. the F-sharp) is further removed from the other notes.

```
# Python 3.10.9
# gensim==4.3.1, numpy==1.24.2, scipy==1.10.1
# scikit-learn==1.2.2, smart-open==6.3.0

from gensim.models.word2vec import Word2Vec
from sklearn.decomposition import PCA
themes = [list('01213014545144'), list('54104516740451')]
w2v = Word2Vec(themes, min_count=1, vector_size=5)
pca = PCA(n_components=2)
X = w2v.wv[w2v.wv.key_to_index]
transformed = pca.fit_transform(X)
result = list(
    (word, transformed[i, 0], transformed[i, 1]) for i, word in (
        enumerate(w2v.wv.key_to_index.keys())
    )
)
```

3.8. Methodological utility

Elements of the methodologies demonstrated in this chapter can therefore be used to conduct analysis for monophonic music as modelled through a natural language processing lens. In the hypothetical scenario where the corpus of ND polyphony were

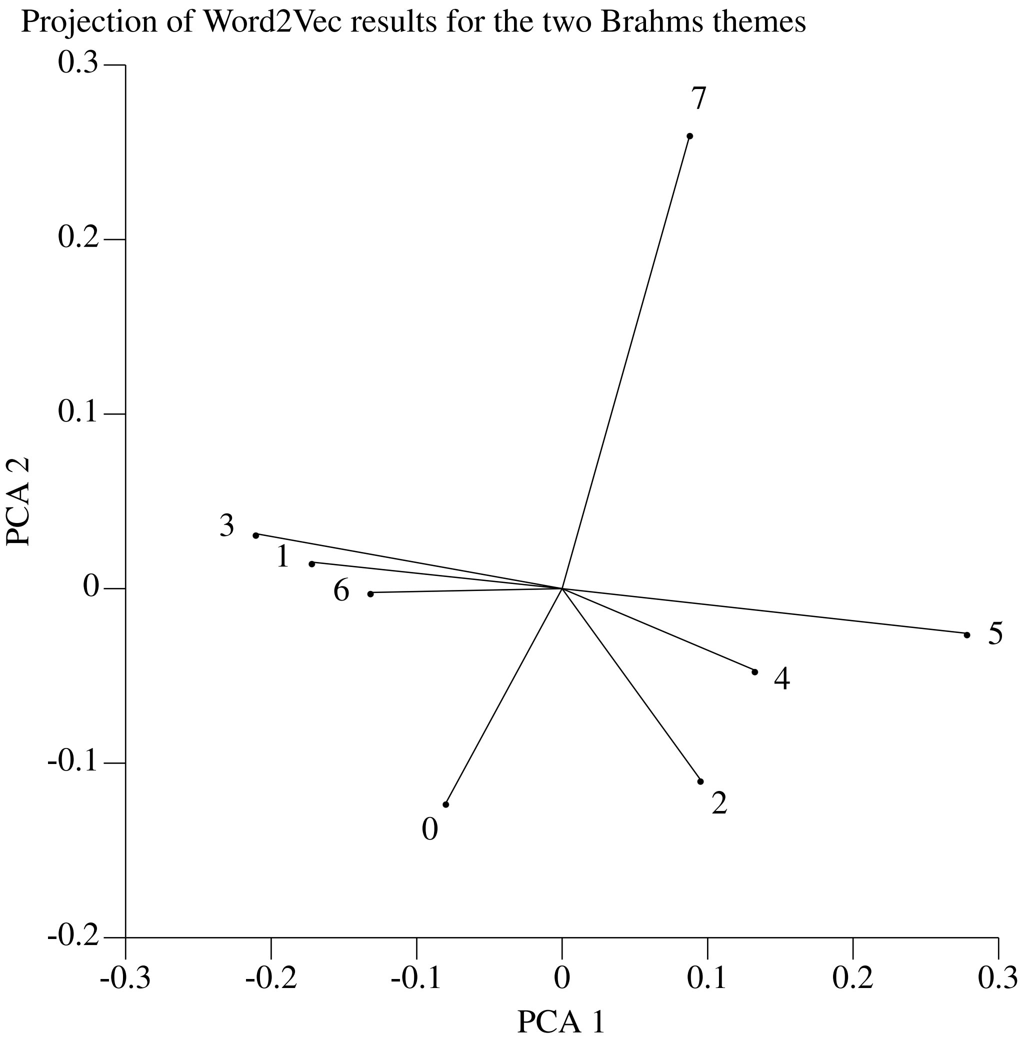

Fig. 3.27

indeed monophonic, how would the initial research question be answered? To reiterate, I am looking for elements of musical reuse: what sequences or musical ideas occur frequently within the corpus? Direct comparisons between settings of music are not that useful. Saying that one piece of music is similar to another does not tell us much about sections (clausulae, indirect concordances, or musical formulae?) within that piece that are perhaps of note. Sequence alignment will therefore not be useful.

Moreover, as the number of pairwise comparisons between sections of each piece increases, the task of cross-referencing similarity scores for every possible combination becomes impractical and unfeasible, and so I must therefore also discount the usefulness of calculating edit distances. If there were a small number of pieces in the corpus that I already knew were alike and wanted to see which were closest — perhaps for stemmatic analysis — then these metrics may have been useful. As it stands however, in this case there are hundreds, if not thousands, of pieces that are similar in small ways within limited sets. Say there were only 100 pieces, then the combinations of pairwise comparisons would be:

$$^{100}C_2 = \frac{100!}{2!(100-2)!} = 4950 \text{ comparisons}$$

If I consider a hypothetical figure of each piece being made of ten sections, then the number of comparisons balloons to half a million. If I try then to be more granular by considering 1000 n-grams in each piece, then it takes it to half a billion comparisons. Such numbers are unfeasibly large to process: even the conservative estimate of half a billion comparisons against an optimistic estimate of 100 comparisons each second would still take two months to process. The same data must then also be processed at different n and k values. Rather than making comparisons one at a time, a methodology that can map out the corpus so that similar sections can be clustered together — such as Word2Vec-like vectorisation — will be much more efficient as each point will only need to be compared to its neighbours.

For a monophonic corpus, then, the ideal methodology would be to encode small sections (such as ever-increasing n-grams or skip-grams) in an embedding, mapping these n-grams into a d-dimensional space, and clustering similar items together. Given our discussion has so far been limited to a monophonic corpus, what remains to be described therefore is a methodology to accomplish the same goal in a polyphonic context. A solution to this will presented in “Chapter 6: Offline”.

3.9. Conclusion

This chapter has aimed to serve three purposes: firstly to contextualise this study within a philosophical framework that addresses the intricacies of data modelling in music; secondly to examine common data structures for encoding music, particularly focusing on the representation of ND polyphony; and finally to explore corpus methodologies that can be adapted to contribute to this study’s methodology.

In the realm of data modelling, it is essential to approach the creation of models with caution, recognising that they inevitably reflect our biases, whether justly or unjustly. Each model that we might create is an abstraction for drawing out relationships between structures that we see in the source material, and each reflects a different selection of our biases. As Bohlman (2001, p.17) eloquently states, ‘music may be what we think it is; it may not be’. If the former holds true, our models face no inherent issues. However, if the latter is the case, all our current models may be flawed. Uncertainty prevails, necessitating a critical stance towards data models that adhere to specific theories of “how music is”. This is particularly important when dealing with the intricate and ambiguous nature of “music” at large. Instead, we should aim to create models that can accommodate current thinking while remaining open to alternative perspectives. These models should practice the separation of concerns, and the flow of data between different musical domains should be meticulously transparent.

When considering the representation of ND polyphony within the logical domain, it becomes evident that neither tabular nor hierarchical data models suffice without imposing biased or excessively editorial constraints. Therefore, this study will adopt a graph data model to represent its logical data as, rather than reflecting relationships based on absolute sequence and alignment, it can represent ambiguous and complex, cross-structural relationships. Although this comes at the expense of human readability and sequential parsing, I believe that this trade-off is worth it because the primary aim of this study is to perform computer analysis, not to create human-readable encodings. Such a model can be constructed using a non-standard MEI dialect, albeit with some flaws yet to be fully identified and solutions worked out.

Lastly, among the corpus methodologies surveyed, the most valuable approaches for the simpler case of monophony involve the utilisation of n-grams and token vectorisation such as CBOW. "Chapter 6: Offline" will demonstrate how this is relevant also to ND notation. While edit distances and sequence alignment may appear to be pertinent, their reliance on pairwise and whole-sequence comparisons renders them impractical for the large scale of this study. However, it is crucial to emphasise that regardless of the chosen methodology, the outcomes are unlikely to yield definitive yes/no answers, such as a comforting "*x* is an example of musical reuse, *y* is not". Instead, the results obtained through these methodologies manifest as probabilities or proportions, and as such rely on the parameters of the chosen model.

Karp (1966) encoded using *Humdrum*

```
!!!OTL: Pascha nostrum
!!!SMS: D-W Cod. Guelf. 628 Helmst.
!!!SML: Herzog August Bibliothek, Wolfenbüttel, Germany
!!!ENC: Joshua Stutter
!!!RWG: Using the non-standard **NotreDame representation
!!!voices: 2
**kern  **NotreDame **kern  **NotreDame **text  **kern  **NotreDame **text
*part2  *part2      *part1L *part1L     *part1L *part1B *part1B     *
*clefC4 *           *clefF3 *           *       *clefF3 *           *
*k[]    *           *k[]    *           *       *k[]    *           *
c       .           F       .           Pa-     .       .           Pa-
.       |           .       |           .       .       |           .
c       [           .       .           .       .       .           .
d       ]           .       .           .       .       .           .
d       v           .       .           .       .       .           .
.       |           .       .           .       .       .           .
d       [           E       .           -scha   .       .           .
e       ]           .       |           .       E       .           -scha
d       [           .       .           .       .       |           .
e       ]           .       .           .       .       .           .
.       |           .       .           .       .       .           .
f       (           .       .           .       .       .           .
e       .           .       .           .       .       .           .
d       .           .       .           .       .       .           .
c       )           .       .           .       .       .           .
B       .           .       .           .       .       .           .
A       [           .       .           .       .       .           .
B       ]           .       .           .       .       .           .
.       |           .       .           .       .       .           .
A       .           .       .           .       .       .           .
B       [           .       .           .       .       .           .
c       .           .       .           .       .       .           .
d       ]           .       .           .       .       .           .
d       [           .       .           .       .       .           .
e       ]           .       .           .       .       .           .
d       (           .       .           .       .       .           .
c       .           .       .           .       .       .           .
B       .           .       .           .       .       .           .
A       )           .       .           .       .       .           .
B       [           .       .           .       .       .           .
c       ]v          .       .           .       .       .           .
.       |           .       .           .       .       .           .
A       [           G       .           no-     .       .           .
G       ]           .       |           .       G       .           no-
.       |           .       .           .       .       |           .
*       *           *clefC4 *           *       *clefC4 *           *
c       .           .       .           .       .       .           .
d       .           .       .           .       .       .           .
d       [           .       .           .       .       .           .
e       .           .       .           .       .       .           .
c       ]           .       .           .       .       .           .
B       [           .       .           .       .       .           .
A       ]           .       .           .       .       .           .
```

Fig. 3.28

Karp (1966) encoded using *Humdrum* (cont.)

```
G       [       .       .       .       .       .       .
c       ]       .       .       .       .       .       .
e       [       .       .       .       .       .       .
d       .       .       .       .       .       .       .
A       ]       A       .       .       A       .       .
.       |       .       |       .       .       |       .
c       .       .       .       .       G       .       .
d       .       .       .       .       .       |       .
d       [       .       .       .       .       .       .
e       .       .       .       .       .       .       .
c       .       .       .       .       .       .       .
B       [       .       .       .       .       .       .
A       ]       .       .       .       .       .       .
G       [       G       .       .       .       .       .
A       ]       .       |       .       .       .       .
B       [       .       .       .       .       .       .
A       .       .       .       .       .       .       .
D       ]       A       .       .       A       .       .
.       |       .       |       .       .       |       .
```

Fig. 3.29

Roesner (1981) encoded using *Humdrum*

```
!!!OTL: In columbe
!!!ENC: Joshua Stutter
!!!RWG: Using the non-standard **NotreDame representation
!!!voices: 2
**kern  **NotreDame **kern  **NotreDame **kern  **NotreDame **kern  **NotreDame **kern  **NotreDame **kern  **NotreDame **text
*pt2W1  *pt2W1      *pt1W1  *pt1W1      *pt2F   *pt2F       *pt1F   *pt1F       *pt2W2  *pt2W2      *pt1W2  *pt1W2      *
*clefC4 *           *clefF3 *           *clefC4 *           *clefC4 *           *clefC4 *           *clefF3 *           *
*k[]    *           *k[]    *           *k[]    *           *k[]    *           *k[]    *           *k[]    *           *
f       (B          F       L.2         f       (L          F       L.2         f       [L          F       L.2         super
e       B           .       .           e       B           .       .           e       B           .       .           .
d       B           .       .           d       L           .       .           d       ]L          .       .           .
c       )L.         .       .           c       )B          .       .           c       [B          .       .           .
e       (B          E       L.2         e       (L          E       L.2         e       ]L          E       L.2         .
d       B           .       .           d       B           .       .           d       [B          .       .           .
c       B           .       .           c       L           .       .           c       ]L          .       .           .
B       )L.         .       .           B       )B          .       .           B       [B          .       .           .
d       (B          D       L.2         d       (L          D       L.6         d       ]L          D       L.2         .
c       B           .       .           c       B           .       .           c       [B          .       .           .
B       B           .       .           B       L           .       .           B       ]L          .       .           .
A       )L.         .       .           A       )B          .       .           A       [B          .       .           .
c       (B          C       L.3         c       [L          .       .           .       .           .       .           .
B       B           .       |           B       ]B          .       .           .       .           .       .           .
A       B           .       .           B       (L          .       .           .       .           .       .           .
G       )L.         .       .           A       B           .       .           .       .           .       .           .
G       L           .       .           G       )L.         .       .           .       .           .       .           .
.       |B          .       .           A       [L          .       .           .       .           .       .           .
.       .           .       .           B       B           .       .           .       .           .       .           .
.       .           .       .           c       L.          C       L.          c       L.          C       L.          .
.       .           .       .           .       |L.         .       |L.         .       |L.         .       |L.         .
d       B           D       L.          d       [L          D       L.2         d       [L          D       L.2         a-
e       [B          .       |           e       B           .       .           e       B           .       .           .
d       ]B          .       .           d       ]LvB        .       .           d       ]LvB        .       .           .
c       L           F       L.          c       [L          F       L.2         c       L[          F       L.2         -quas
```

Fig. 3.30

Roesner (1981) encoded using *Humdrum* (cont.)

```
.	|B	.	|	d	B	.	.	d	B	.	.	.
.	.	.	.	c	]L	.	.	c	L]	.	.	.
.	.	.	.	A	[B	.	.	A	B[	.	.	.
e	[L	E	L.	B	]L	E	L.2	B	L]	E	L.2	.
d	B	.	.	c	[B	.	.	c	B[	.	.	.
.	.	.	.	B	]L	.	.	B	L]	.	.	.
.	.	.	.	G	[B	.	.	G	B[	.	.	.
c	](BvB	F	L.	A	]L	F	L.2	A	L]	F	L.2	.
.	.	.	.	B	[B	.	.	B	B[	.	.	.
A	B	.	.	A	]L.	.	.	A	]L.	.	.	.
.	.	.	.	.	.	.	.	.	|	.	.	.
G	)B	G	L.	G	L.	G	L.2	G	L.	G	L.2	.
A	[B	.	.	A	[L	.	.	A	L[	.	.	.
B	B	.	.	B	B	.	.	B	B	.	.	.
c	]L	F	L.	c	]L.	F	L.	c	]L.	F	L.	.
.	|B	.	|	.	|L.	.	|L.	.	|L.	.	|L.	.
f	(B	F	L.2	f	(B	F	L.2	f	(B	F	L.2	mul-
e	B	.	.	e	B	.	.	e	B	.	.	.
d	B	.	.	d	B	.	.	d	B	.	.	.
c	)L.	.	.	c	)L.	.	.	c	)L.	.	.	.
B	BvB	E	L.2	e	(B	E	L.2	e	(B	E	L.2	.
.	.	.	.	d	B	.	.	d	B	.	.	.
B	[B	.	.	c	B	.	.	c	B	.	.	.
d	]L.	.	.	B	)L.	.	.	B	)L.	.	.	.
d	L.	D	L.2	d	[B	D	L.2	d	[B	D	L.2	.
e	[L	.	.	e	B	.	.	e	B	.	.	.
d	B	.	.	d	B	.	.	d	B	.	.	.
c	]L	F	L.3	c	]L.	.	.	c	]L.	.	.	.
.	|B	.	|	.	.	.	.	.	.	.	.	.
A	L	.	.	.	.	.	.	.	.	.	.	.
B	[B	.	.	.	.	.	.	.	.	.	.	.
c	]L	.	.	c	L	F	L.	c	L	F	L.	.
.	|B	.	.	.	|B	.	|	.	|B	.	|	.
```

Fig. 3.31

Karp (1966) encoded using MEI

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<?xml-model
  href="http://example.com/schema/4.0.0/mei-notredame.rng"
  type="application/xml"
  schematypens="http://relaxng.org/ns/structure/1.0"
?>
<mei xmlns="http://example.com/ns/mei-notredame">
  <meiHead>
    <fileDesc>
      <titleStmt>
        <title>Pascha nostrum</title>
        <respStmt>
          <persName role="encoder">Joshua Stutter</persName>
        </respStmt>
      </titleStmt>
      <sourceDesc>
        <source>
          <bibl>
            <identifier type="URI">
             http://diglib.hab.de/mss/628-helmst/start.htm?image=00067
            </identifier>
          </bibl>
        </source>
        <source xml:id="ludwig">Friedrich Ludwig</source>
        <source xml:id="bukofzer">Manfred Bukofzer</source>
      </sourceDesc>
    </fileDesc>
  </meiHead>
  <music>
    <body>
      <mdiv>
        <score>
          <scoreDef>
            <staffGrp>
              <staffDef n="1" clef.line="4" clef.shape="C" lines="5" />
              <staffDef n="2" clef.line="3" clef.shape="F" lines="4" />
            </staffGrp>
          </scoreDef>
          <section>
            <staff n="1">
              <layer>
                <note oct="4" pname="c" xml:id="pa1" />
                <divisione />
                <ligature type="square">
                  <note oct="4" pname="c" />
                  <note oct="4" pname="d" />
                </ligature>
                <note oct="4" pname="d" plica="down" />
                <divisione />
                <ligature type="square">
                  <note oct="4" pname="d" xml:id="scha1" />
                  <note oct="4" pname="e" xml:id="scha2" />
                </ligature>
```

Fig. 3.32

Karp (1966) encoded using MEI (cont.)

```
          <ligature type="square">
            <note oct="4" pname="d" />
            <note oct="4" pname="e" />
          </ligature>
          <divisione />
          <ligature type="currentes">
            <note oct="4" pname="f" />
            <note oct="4" pname="e" />
            <note oct="4" pname="d" />
            <note oct="4" pname="c" />
          </ligature>
          <note oct="3" pname="b" />
          <ligature type="square">
            <note oct="3" pname="a" />
            <note oct="3" pname="b" />
          </ligature>
          <divisione />
          <note oct="3" pname="a" />
          <ligature type="square">
            <note oct="3" pname="b" />
            <note oct="4" pname="c" />
            <note oct="4" pname="d" />
          </ligature>
          <ligature type="square">
            <note oct="4" pname="d" />
            <note oct="4" pname="e" />
          </ligature>
          <ligature type="currentes">
            <note oct="4" pname="d" />
            <note oct="4" pname="c" />
            <note oct="3" pname="b" />
            <note oct="3" pname="a" />
          </ligature>
          <ligature type="square">
            <note oct="3" pname="b" />
            <note oct="4" pname="c" plica="down" />
          </ligature>
          <divisione />
          <ligature type="square">
            <note oct="3" pname="a" xml:id="no1" />
            <note oct="3" pname="g" xml:id="no2" />
          </ligature>
          <divisione />
        </layer>
      </staff>
      <staff n="2">
        <layer>
          <note oct="3" pname="f" syl="Pas-" synch="#pa1" />
          <divisione />
          <app>
            <rdg source="#ludwig">
              <note oct="3" pname="e" syl="-scha" synch="#scha1" />
            </rdg>
```

Fig. 3.33

Karp (1966) encoded using MEI (cont.)

```
            <rdg source="#bukofzer">
              <note oct="3" pname="e" syl="-scha" synch="#scha2" />
            </rdg>
          </app>
          <divisione />
          <app>
            <rdg source="#ludwig">
              <note oct="3" pname="g" syl="no-" synch="#no1" />
            </rdg>
            <rdg source="#bukofzer">
              <note oct="3" pname="g" syl="no-" synch="#no2" />
            </rdg>
          </app>
        </layer>
      </staff>
    </section>
    <section>
      <staff n="1">
        <layer>
          <note oct="4" pname="c" />
          <note oct="4" pname="d" />
          <ligature type="square">
            <note oct="4" pname="d" />
            <note oct="4" pname="e" />
            <note oct="4" pname="c" />
          </ligature>
          <ligature type="square">
            <note oct="3" pname="b" />
            <note oct="3" pname="a" />
          </ligature>
          <ligature type="square">
            <note oct="3" pname="g" />
            <note oct="4" pname="c" />
          </ligature>
          <ligature type="square">
            <note oct="4" pname="e" />
            <note oct="4" pname="d" />
            <note oct="3" pname="a" xml:id="no3" />
          </ligature>
          <divisione />
          <note oct="4" pname="c" xml:id="no4" />
          <note oct="4" pname="d" />
          <ligature type="square">
            <note oct="4" pname="d" />
            <note oct="4" pname="e" />
            <note oct="4" pname="c" />
          </ligature>
          <ligature type="square">
            <note oct="3" pname="b" />
            <note oct="3" pname="a" />
          </ligature>
          <ligature type="square">
            <note oct="3" pname="g" xml:id="no5" />
```

Fig. 3.34

Karp (1966) encoded using MEI (cont.)

```
                <note oct="3" pname="a" />
              </ligature>
              <ligature type="square">
                <note oct="3" pname="b" />
                <note oct="3" pname="a" />
                <note oct="3" pname="d" xml:id="no6" />
              </ligature>
              <divisione />
            </layer>
          </staff>
          <staff n="2">
            <clef line="4" shape="C" lines="4" />
            <note oct="3" pname="a" synch="#no3" />
            <divisione />
            <app>
              <rdg source="#ludwig">
                <note oct="3" pname="g" synch="#no4" />
              </rdg>
              <rdg source="#bukofzer">
                <note oct="3" pname="g" synch="#no5" />
              </rdg>
            </app>
            <note oct="3" pname="a" synch="#no6" />
          </staff>
        </section>
      </score>
    </mdiv>
  </body>
</music>
</mei>
```

Fig. 3.35

Roesner (1981) encoded using MEI

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<?xml-model
  href="http://example.com/schema/4.0.0/mei-notredame.rng"
  type="application/xml"
  schematypens="http://relaxng.org/ns/structure/1.0"?>
<mei xmlns="http://example.com/ns/mei-notredame">
  <meiHead>
    <fileDesc>
      <titleStmt>
        <title>In columbe</title>
        <respStmt>
          <persName role="encoder">Joshua Stutter</persName>
        </respStmt>
      </titleStmt>
      <sourceDesc>
        <source xml:id="w1">
          <bibl>
            <identifier type="URI">
              http://diglib.hab.de/mss/628-helmst/start.htm?image=00039
            </identifier>
            <physLoc>
              <identifier
               type="RISM"
               auth="DIAMM"
               auth.uri="https://www.diamm.ac.uk/sources/870/#/"
              >
                D-W Cod. Guelf. 628 Helmst.
              </identifier>
              <repository>
                <country>Germany</country>
                <settlement>Wolfenbüttel</settlement>
                <name>Herzog August Bibliothek</name>
              </repository>
            </physLoc>
          </bibl>
        </source>
        <source xml:id="w2">
          <bibl>
            <identifier type="URI">
              http://diglib.hab.de/mss/1099-helmst/start.htm?image=00107
            </identifier>
            <physLoc>
              <identifier
               type="RISM"
               auth="DIAMM"
               auth.uri="https://www.diamm.ac.uk/sources/854/#/"
              >
                D-W Cod. Guelf. 1099 Helmst.
              </identifier>
              <repository>
                <country>Germany</country>
                <settlement>Wolfenbüttel</settlement>
                <name>Herzog August Bibliothek</name>
```

Fig. 3.36

Roesner (1981) encoded using MEI (cont.)

```
              </repository>
            </physLoc>
          </bibl>
        </source>
        <source xml:id="f">
          <bibl>
            <identifier type="URI">
              http://mss.bmlonline.it/s.aspx?Id=AWOHy_N-
I1A4r7GxMB57&c=Antiphonarium#/oro/112
            </identifier>
            <physLoc>
              <identifier
               type="RISM"
               auth="DIAMM"
               auth.uri="https://www.diamm.ac.uk/sources/924/#/"
              >
                I-Fl MS Pluteus 29.1
              </identifier>
              <repository>
                <country>Italy</country>
                <settlement>Firenze</settlement>
                <name>Biblioteca Medicea-Laurenziana</name>
              </repository>
            </physLoc>
          </bibl>
        </source>
      </sourceDesc>
    </fileDesc>
  </meiHead>
  <music>
    <body>
      <mdiv>
        <score>
          <scoreDef>
            <staffGrp>
            <staffDef n="1" lines="5" />
            <staffDef n="2" lines="4" />
            </staffDef>
          </scoreDef>
          <section>
            <staff n="1">
              <layer>
                <clef line="4" type="c" />
                <app>
                <rdg source="#w1">
                  <ligature type="currentes" xml:id="w1s1e1" corresp="#fs1e1">
                    <note oct="4" pname="f" xml:id="w1s1e1-1" num="2" numbase="1"
                      synch="#s2e1" corresp="#fs1e1-1 #w2s1e1-1" />
                    <note oct="4" pname="e" xml:id="w1s1e1-2" num="2" numbase="1"
                      corresp="#fs1e1-2 #w2s1e1-2" />
                    <note oct="4" pname="d" xml:id="w1s1e1-3" num="2" numbase="1"
                      corresp="#fs1e1-3 #w2s1e1-3" />
                    <note oct="4" pname="c" xml:id="w1s1e1-4" num="6" numbase="1"
```

Fig. 3.37

Roesner (1981) encoded using MEI (cont.)

```
        corresp="#fs1e1-4 #w2s1e2-1" />
    </ligature>
    <ligature type="currentes" xml:id="w1s1e2" corresp="#fs1e2">
      <note oct="4" pname="e" xml:id="w1s1e2-1" num="2" numbase="1"
        synch="#s2e2" corresp="#fs1e2-1 #w2s1e2-2" />
      <note oct="4" pname="d" xml:id="w1s1e2-2" num="2" numbase="1"
        corresp="#fs1e2-2 #w2s1e3-1" />
      <note oct="4" pname="c" xml:id="w1s1e2-3" num="2" numbase="1"
        corresp="#fs1e2-3 #w2s1e3-2" />
      <note oct="3" pname="b" xml:id="w1s1e2-4" num="6" numbase="1"
        corresp="#fs1e2-4 #w2s1e4-1" />
    </ligature>
    <ligature type="currentes" xml:id="w1s1e3" corresp="#fs1e3">
      <note oct="4" pname="d" xml:id="w1s1e3-1" num="2" numbase="1"
        synch="#w1w2s2e2" corresp="#fs1e3-1 #w2s1e4-2" />
      <note oct="4" pname="c" xml:id="w1s1e3-2" num="2" numbase="1"
        corresp="#fs1e3-2 #w2s1e5-1" />
      <note oct="3" pname="b" xml:id="w1s1e3-3" num="2" numbase="1"
        corresp="#fs1e3-3 #w2s1e5-2" />
      <note oct="3" pname="a" xml:id="w1s1e3-4" num="6" numbase="1"
        corresp="#fs1e3-4 #w2s1e6-1" />
    </ligature>
    <ligature type="currentes" xml:id="w1s1e4">
      <note oct="4" pname="c" xml:id="w1s1e4-1" num="2" numbase="1"
        synch="#w1s2e3" corresp="#fs1e4-1" />
      <note oct="3" pname="b" xml:id="w1s1e4-2" num="2" numbase="1"
        corresp="#fs1e4-2" />
      <note oct="3" pname="a" xml:id="w1s1e4-3" num="2"
        numbase="1" />
      <note oct="3" pname="g" xml:id="w1s1e4-4" num="6"
        numbase="1" />
    </ligature>
    <note oct="3" pname="3" xml:id="w1s1e5" num="4" numbase="1" />
    <divisione num="2" numbase="1" xml:id="w1s1e6"
      corresp="#fs1e7 #w2s1e7" />
  </rdg>
  <rdg source="#f">
    <ligature type="currentes" xml:id="fs1e1" corresp="#w1s1e1">
      <note oct="4" pname="f" xml:id="fs1e1-1" num="4" numbase="1"
        synch="#s2e1" corresp="#w1s1e1-1 #w2s1e1-1" />
      <note oct="4" pname="e" xml:id="fs1e1-2" num="2" numbase="1"
        corresp="#w1s1e1-2 #w2s1e1-2" />
      <note oct="4" pname="d" xml:id="fs1e1-3" num="4" numbase="1"
        corresp="#w1s1e1-3 #w2s1e1-3" />
      <note oct="4" pname="c" xml:id="fs1e1-4" num="2" numbase="1"
        corresp="#w1s1e1-4 #w2s1e2-1" />
    </ligature>
    <ligature type="currentes" xml:id="fs1e2" corresp="#w1s1e2">
      <note oct="4" pname="e" xml:id="fs1e2-1" num="4" numbase="1"
        synch="#s2e2" corresp="#w1s1e2-1 #w2s1e2-2" />
      <note oct="4" pname="d" xml:id="fs1e2-2" num="2" numbase="1"
        corresp="#w1s1e2-2 #w2s1e3-1" />
      <note oct="4" pname="c" xml:id="fs1e2-3" num="4" numbase="1"
```

Fig. 3.38

Roesner (1981) encoded using MEI (cont.)

```
        corresp="#w1s1e2-3 #w2s1e3-2" />
      <note oct="3" pname="b" xml:id="fs1e2-4" num="2" numbase="1"
        corresp="#w1s1e2-4 #w2s1e4-1" />
    </ligature>
    <ligature type="currentes" xml:id="fs1e3" corresp="#w1s1e3">
      <note oct="4" pname="d" xml:id="fs1e3-1" num="4" numbase="1"
        synch="#fs2e2" corresp="#w1s1e3-1 #w2s1e4-2" />
      <note oct="4" pname="c" xml:id="fs1e3-2" num="2" numbase="1"
        corresp="#w1s1e3-2 #w2s1e5-1" />
      <note oct="3" pname="b" xml:id="fs1e3-3" num="4" numbase="1"
        corresp="#w1s1e3-3 #w2s1e5-2" />
      <note oct="3" pname="a" xml:id="fs1e3-4" num="2" numbase="1"
        corresp="#w1s1e3-4 #w2s1e6-1" />
    </ligature>
    <ligature type="square" xml:id="fs1e4">
      <note oct="4" pname="c" xml:id="fs1e4-1" num="4" numbase="1"
        corresp="#w1s1e4-1" />
      <note oct="3" pname="b" xml:id="fs1e4-2" num="2" numbase="1"
        coresp="#w1s1e4-2" />
    </ligature>
    <ligature type="currentes" xml:id="fs1e5">
      <note oct="3" pname="b" xml:id="fs1e5-1" num="4"
        numbase="1" />
      <note oct="3" pname="a" xml:id="fs1e5-2" num="2"
        numbase="1" />
      <note oct="3" pname="g" xml:id="fs1e5-3" num="6"
        numbase="1" />
    </ligature>
    <ligature type="square" xml:id="fs1e6">
      <note oct="3" pname="a" xml:id="fs1e6-1" num="4"
        numbase="1" />
      <note oct="3" pname="b" xml:id="fs1e6-2" num="2"
        numbase="1" />
      <note oct="4" pname="c" xml:id="fs1e6-3" num="6"
        numbase="1" synch="#fw2s2e3" corresp="#w2s1e6-2" />
    </ligature>
    <divisione num="6" numbase="1" xml:id="fs1e7"
      corresp="#w1s1e6 #w2s1e7" />
  </rdg>
  <rdg source="#w2">
  <ligature type="square" xml:id="w2s1e1">
    <note oct="4" pname="f" xml:id="w2s1e1-1" num="4" numbase="1"
      synch="#s2e1" corresp="#w1s1e1-1 #fs1e1-1" />
    <note oct="4" pname="e" xml:id="w2s1e1-2" num="2" numbase="1"
      corresp="#w1s1e1-2 #fs1e1-2" />
    <note oct="4" pname="d" xml:id="w2s1e1-3" num="4" numbase="1"
      corresp="#w1s1e1-3 #fs1e1-3" />
  </ligature>
  <ligature type="square" xml:id="w2s1e2">
    <note oct="4" pname="c" xml:id="w2s1e2-1" num="2" numbase="1"
      corresp="#w1s1e1-4 #fs1e1-4" />
    <note oct="4" pname="e" xml:id="w2s1e2-2" num="4" numbase="1"
      synch="#s2e2" corresp="#w1s1e2-1 #fs1e2-1" />
```

Fig. 3.39

Roesner (1981) encoded using MEI (cont.)

```
      </ligature>
      <ligature type="square" xml:id="w2s1e3">
        <note oct="4" pname="d" xml:id="w2s1e3-1" num="2" numbase="1"
          corresp="#w1s1e2-2 #fs1e2-2" />
        <note oct="4" pname="c" xml:id="w2s1e3-2" num="4" numbase="1"
          corresp="#w1s1e2-3 #fs1e2-3" />
      </ligature>
      <ligature type="square" xml:id="w2s1e4">
        <note oct="3" pname="b" xml:id="w2s1e4-1" num="2" numbase="1"
          corresp="#w1s1e2-4 #fs1e2-4" />
        <note oct="4" pname="d" xml:id="w2s1e4-2" num="4" numbase="1"
          synch="#w1w2s2e2" corresp="#w1s1e3-1 #fs1e3-1" />
      </ligature>
      <ligature type="square" xml:id="w2s1e5">
        <note oct="4" pname="c" xml:id="w2s1e5-1" num="2" numbase="1"
          corresp="#w1s1e3-2 #fs1e3-2" />
        <note oct="3" pname="b" xml:id="w2s1e5-2" num="4" numbase="1"
          corresp="#w1s1e3-3 #fs1e3-3" />
      </ligature>
      <ligature type="square" xml:id="w2s1e6">
        <note oct="3" pname="a" xml:id="w2s1e6-1" num="2" numbase="1"
          corresp="#w1s1e3-4 #fs1e3-4" />
        <note oct="4" pname="c" xml:id="w2s1e6-2" num="6" numbase="1"
          synch="#fw2s2e3" corresp="#fs1e6-3" />
      </ligature>
      <divisione num="6" numbase="1" xml:id="w2s1e7"
        corresp="#w1s1e6 #fs1e7" />
    </layer>
  </staff>
  <staff n="2">
    <layer>
      <app>
        <rdg source="#w1 #w2">
          <clef line="3" shape="F" />
        </rdg>
        <rdg source="#f">
          <clef line="4" shape="C" />
        </rdg>
      </app>
      <note oct="3" pname="f" num="12" numbase="1" xml:id="s2e1"
        synch="#w1s1e1-1 #fs1e1-1 #w2s1e1-1" syl="super" />
      <note oct="3" pname="e" num="12" numbase="1" xml:id="s2e2"
        synch="#w1s1e2-1 #fs1e2-1 #w2s1e2-2" />
      <app>
        <rdg source="#w1 #w2">
          <note oct="3" pname="d" num="12" numbase="1"
            xml:id="w1w2s2e2" synch="#w1s1e3-1 #w2s1e4-2"
            corresp="#fs2e2" />
        </rdg>
        <rdg source="#f">
          <note oct="3" pname="d" num="36" numbase="1" xml:id="fs2e2"
            synch="#fs1e3-1" corresp="#w1w2s2e2" />
        </rdg>
```

Fig. 3.40

Roesner (1981) encoded using MEI (cont.)

```
          </app>
          <app>
            <rdg source="#w1">
              <note oct="3" pname="c" num="16" numbase="1" xml:id="w1s2e3"
                synch="#w1s1e4-1" corresp="#fw2s2e3" />
              <divisione num="2" numbase="1" xml:id="w1s2e4"
                corresp="#fw2s2e4" />
            </rdg>
            <rdg source="#f #w2">
              <note oct="3" pname="c" num="6" numbase="1" xml:id="fw2s2e3"
                synch="#fs1e6-3 #w2s1e6-2" corresp="#w1s2e3" />
              <divisione num="6" numbase="1" xml:id="fw2s2e4"
                corresp="#w1s2e4" />
            </rdg>
          </app>
        </layer>
      </staff>
    </section>
    <section>
    <staff n="1">
      <layer>
        <app>
          <rdg source="#w1">
            <note oct="4" pname="d" num="2" numbase="1" xml:id="w1s1e7"
              synch="#w1s2e5" corresp="#fw2s1e8-1" />
            <ligature type="square" xml:id="w1s1e8">
              <note oct="4" pname="e" num="2" numbase="1" xml:id="w1s1e8-1"
                corresp="#fw2s1e8-2" />
              <note oct="4" pname="d" num="2" numbase="1" xml:id="w1s1e8-2"
                corresp="#fw2s1e8-3" />
            </ligature>
            <note oct="4" pname="c" num="4" numbase="1" xml:id="w1s1e9"
              synch="#w1s2e7" corresp="#fw2s1e9-1" />
            <divisione num="2" numbase="1" xml:id="w1s1e10"
              synch="#w1s2e8" />
            <ligature type="square" xml:id="w1s1e11">
              <note oct="4" pname="e" num="4" numbase="1"
                xml:id="w1s1e11-1" synch="#w1s2e9" />
              <note oct="4" pname="d" num="2" numbase="1"
                xml:id="w1s1e11-2" />
              <note oct="4" pname="c" num="4" numbase="1" plica="down"
                xml:id="w1s1e11-3" synch="#w1s2e10" />
            </ligature>
            <ligature type="currentes" xml:id="w1s1e12">
              <note oct="3" pname="a" num="2" numbase="1"
                xml:id="w1s1e12-1" />
              <note oct="3" pname="g" num="2" numbase="1"
                xml:id="w1s1e12-2" synch="#w1s2e11" />
            </ligature>
            <ligature type="square" xml:id="w1s1e13" corresp="#fw2s1e14">
              <note oct="3" pname="a" num="2" numbase="1"
                xml:id="w1s1e13-1" corresp="#fw2s1e14-1" />
              <note oct="3" pname="b" num="2" numbase="1"
```

Fig. 3.41

Roesner (1981) encoded using MEI (cont.)

```
        xml:id="w1s1e13-2" corresp="#fw2s1e14-2" />
      <note oct="4" pname="c" num="4" numbase="1"
        xml:id="w1s1e13-3" corresp="#fw2s1e14-3" />
    </ligature>
    <divisione num="2" numbase="1" xml:id="w1s1e14"
      synch="#fw2s2e11" corresp="#fw2s1e15" />
  </rdg>
  <rdg source="#f #w2">
    <ligature type="square" xml:id="fw2s1e8">
      <note oct="4" pname="d" num="4" numbase="1"
        synch="#w1s2e5" xml:id="fw2s1e8-1"
        corresp="#w1s1e7" />
      <note oct="4" pname="e" num="2" numbase="1"
        xml:id="fw2s1e8-2" corresp="#w1s1e8-1" />
      <note oct="4" pname="d" num="4" numbase="1" plica="down"
        xml:id="fw2s1e8-3" corresp="#w1s1e8-2" />
    </ligature>
    <ligature type="square" xml:id="fw2s1e9">
      <note oct="4" pname="c" num="4" numbase="1"
        synch="#w1s2e7" xml:id="fw2s1e9-1"
        corresp="#w1s1e9" />
      <note oct="4" pname="d" num="2" numbase="1"
        xml:id="fw2s1e9-2" />
      <note oct="4" pname="c" num="4" numbase="1"
        xml:id="fw2s1e9-3" />
    </ligature>
    <ligature type="square" xml:id="fw2s1e10">
      <note oct="3" pname="a" num="2" numbase="1"
        xml:id="fw2s1e10-1" />
      <note oct="3" pname="b" num="4" numbase="1"
        xml:id="fw2s1e10-2" synch="#w1s2e9 "/>
    </ligature>
    <ligature type="square" xml:id="fw2s1e11">
      <note oct="4" pname="c" num="2" numbase="1"
        xml:id="fw2s1e11-1" />
      <note oct="3" pname="b" num="4" numbase="1"
        xml:id="fw2s1e11-2" />
    </ligature>
    <ligature type="square" xml:id="fw2s1e12">
      <note oct="3" pname="g" num="2" numbase="1"
        xml:id="fw2s1e12-1" />
      <note oct="3" pname="a" num="4" numbase="1"
        xml:id="fw2s1e12-2" synch="#w1s2e10 "/>
    </ligature>
    <ligature type="square" xml:id="fw2s1e13">
      <note oct="3" pname="b" num="2" numbase="1"
        xml:id="fw2s1e13-1" />
      <note oct="3" pname="a" num="6" numbase="1"
        xml:id="fw2s1e13-2" />
    </ligature>
    <app>
      <rdg source="#f"></rdg>
      <rdg source="#w2">
```

Fig. 3.42

Roesner (1981) encoded using MEI (cont.)

```
                <divisione />
              </rdg>
            </app>
            <note oct="3" pname="g" num="6" numbase="1"
              xml:id="fw2s1e14" synch="#w1s2e11" />
            <ligature type="square" xml:id="fw2s1e14" corresp="#w1s1e13">
              <note oct="3" pname="a" num="4" numbase="1"
                xml:id="fw2s1e14-1" corresp="#w1s1e13-1" />
              <note oct="3" pname="b" num="2" numbase="1"
                xml:id="fw2s1e14-2" corresp="#w1s1e13-2" />
              <note oct="4" pname="c" num="6" numbase="1"
                xml:id="fw2s1e14-3" corresp="#w1s1e13-3"
                synch="#w1s2e12" />
            </ligature>
            <divisione num="6" numbase="1" xml:id="fw2s1e15"
              synch="#fw2s2e11" corresp="#w1s1e14" />
          </rdg>
        </app>
      </layer>
    </staff>
    <staff n="2">
      <layer>
        <app>
          <rdg source="#w1">
            <note oct="3" pname="d" num="4" numbase="1" xml:id="w1s2e5"
              syl="a-" corresp="#fw2s2e5" synch="#w1s1e7 #fw2s1e8-1" />
            <divisione num="2" numbase="1" xml:id="w1s2e6" />
            <note oct="3" pname="f" num="4" numbase="1" xml:id="w1s2e7"
              syl="-quas" corresp="#fw2s2e6" synch="#w1s1e9 #fw2s1e9-1" />
            <divisione num="2" numbase="1" xml:id="w1s2e8"
              synch="#w1s1e10" />
            <note oct="3" pname="e" num="6" numbase="1" xml:id="w1s2e9"
              corresp="#fw2s2e7" synch="#w1s1e11-1 #fw2s1e10-2" />
            <note oct="3" pname="f" num="6" numbase="1" xml:id="w1s2e10"
              corresp="#fw2s2e8" synch="#w1s1e11-3 #fw2s1e12-2" />
            <note oct="3" pname="g" num="6" numbase="1" xml:id="w1s2e11"
              corresp="#fw2s2e9" synch="#w1s1e12-2 #fw2s1e14" />
            <note oct="3" pname="f" num="4" numbase="1" xml:id="w1s2e12"
              corresp="#fw2s2e10" synch="#w1s1e13-3 #fw2s1e14-3" />
            <divisione num="2" numbase="1" xml:id="w1s2e13"
              corresp="#fw2s2e11" synch="#w1s1e14 #fw2s1e15" />
          </rdg>
          <rdg source="#f w2">
            <note oct="3" pname="d" num="12" numbase="1" xml:id="fw2s2e5"
              syl="a-" corresp="#w1s2e5" synch="#w1s1e7 #fw2s1e8-1" />
            <note oct="3" pname="f" num="12" numbase="1" xml:id="fw2s2e6"
              syl="-quas" corresp="#w1s2e7" synch="#w1s1e9 #fw2s1e9-1" />
            <note oct="3" pname="e" num="12" numbase="1" xml:id="fw2s2e7"
              corresp="#w1s2e9" synch="#w1s1e11-1 #fw2s1e10-2" />
            <note oct="3" pname="f" num="12" numbase="1" xml:id="fw2s2e8"
              corresp="#w1s2e10" synch="#w1s1e11-3 #fw2s1e12-2" />
            <note oct="3" pname="g" num="12" numbase="1" xml:id="fw2s2e9"
              corresp="#w1s2e11" synch="#w1s1e12-2 #fw2s1e14" />
```

Fig. 3.43

Roesner (1981) encoded using MEI (cont.)

```
                <note oct="3" pname="f" num="6" numbase="1" xml:id="fw2s2e10"
                  corresp="#w1s2e12" synch="#s1s1e13-3 #fw2s1e14-3" />
                <divisione num="6" numbase="1" xml:id="fw2s2e11"
                  corresp="#w1s2e13" synch="#w1s1e14 #fw2s1e15" />
              </rdg>
            </app>
          </layer>
        </staff>
        <section>
          <staff n="1">
            <layer>
              <ligature type="currentes" xml:id="s1e16">
                <note oct="4" pname="f" xml:id="s1e16-1" num="2" numbase="1"
                 synch="#s2e14" />
                <note oct="4" pname="e" xml:id="s1e16-2" num="2" numbase="1" />
                <note oct="4" pname="d" xml:id="s1e16-3" num="2" numbase="1" />
                <note oct="4" pname="c" xml:id="s1e16-4" num="6" numbase="1" />
              </ligature>
              <app>
                <rdg source="#w1">
                  <note oct="3" pname="b" xml:id="w1s1e17" num="4" numbase="1"
                    plica="down" />
                  <ligature type="square" xml:id="w1s1e18">
                    <note oct="3" pname="b" xml:id="w1s1e18-1" num="2"
                      numbase="1" synch="#s2e15" />
                    <note oct="4" pname="d" xml:id="w1s1e18-2" num="6"
                      numbase="1" />
                  </ligature>
                  <note oct="4" pname="d" xml:id="w1s1e19" num="6" numbase="1"
                    corresp="#fw2e18-1" synch="#s2e16" />
                  <ligature type="square" xml:id="w1s1e20">
                    <note oct="4" pname="e" xml:id="w1s1e20-1" num="4"
                      numbase="1" corresp="#fw2e18-2" />
                    <note oct="4" pname="d" xml:id="w1s1e20-2" num="2"
                      numbase="1" corresp="#fw2e18-3" />
                    <note oct="4" pname="c" xml:id="w1s1e20-3" num="4"
                      numbase="1" corresp="#fw2e18-4" synch="#w1s2e17" />
                  </ligature>
                  <divisione num="2" numbase="1" xml:id="w1s1e21" />
                  <note oct="3" pname="a" xml:id="w1s1e22" num="4"
                    numbase="1" />
                  <ligature type="square" xml:id="w1s1e23">
                    <note oct="3" pname="b" xml:id="w1w1e23-1" num="2"
                      numbase="1" />
                    <note oct="4" pname="c" xml:id="w1s1e23-2" num="4"
                      numbase="1" corresp="#fw2e19" />
                  </ligature>
                </rdg>
                <rdg source="#f #w2">
                  <ligature type="currentes" xml:id="fw2e17">
                    <note oct="4" pname="e" xml:id="fw2e17-1" num="2"
                      numbase="1" synch="#s2e15" />
                    <note oct="4" pname="d" xml:id="fw2e17-2" num="2"
```

Fig. 3.44

Roesner (1981) encoded using MEI (cont.)

```
                numbase="1" />
              <note oct="4" pname="c" xml:id="fw2e17-3" num="2"
                numbase="1" />
              <note oct="3" pname="b" xml:id="fw2e17-4" num="6"
                numbase="1" />
            </ligature>
            <ligature type="square" xml:id="fw2e18">
              <note oct="4" pname="d" xml:id="fw2e18-1" num="2"
                numbase="1" corresp="#w1s1e19" synch="#s2e16"/>
              <note oct="4" pname="e" xml:id="fw2e18-2" num="2"
                numbase="1" corresp="#w1s1e20-1" />
              <note oct="4" pname="d" xml:id="fw2e18-3" num="2"
                numbase="1" corresp="#w1s1e20-2" />
              <note oct="4" pname="c" xml:id="fw2e18-4" num="6"
                numbase="1" corresp="#w1s1e20-3" />
            </ligature>
            <note oct="4" pname="c" xml:id="fw2e19" num="4" numbase="1"
              corresp="#w1s1e23-2" synch="#fw2s2e17" />
          </rdg>
        </app>
        <divisione num="2" numbase="1" xml:id="s1e24" synch="#s2e18" />
      </layer>
    </staff>
    <staff n="2">
      <layer>
        <note oct="3" pname="f" xml:id="s2e14" num="12" numbase="1"
          syl="mul-" synch="#s1e16-1" />
        <note oct="3" pname="e" xml:id="s2e15" num="12" numbase="1"
          synch="#w1s1e18-1 #fw2e17-1" />
        <note oct="3" pname="d" xml:id="s2e16" num="12" numbase="1"
          synch="#w1s1e19 #fw2e18-1" />
        <app>
          <rdg source="#w1">
            <note oct="3" pname="f" xml:id="w1s2e17" num="16" numbase="1"
              corresp="#fw2s2e17" synch="#w1s1e20-3" />
          </rdg>
          <rdg source="#f #w2">
            <note oct="3" pname="f" xml:id="fw2s2e17" num="4" numbase="1"
              corresp="#w1s2e17" synch="#fw2e19" />
          </rdg>
        </app>
        <divisione xml:id="s2e18" num="2" numbase="1" synch="#s1e24" />
      </layer>
    </staff>
  </section>
</score>
</mdiv>
</body>
</music>
</mei>
```

Fig. 3.45

Chapter 4
Requirements

The first part of this thesis was concerned primarily with the theory, literature and discussion of the research questions. In dissecting the research questions "What does musical reuse look like beyond the clausula?" and "How can the digital humanities help?" it came to three broad conclusions. In the first chapter it was identified that the sources of ND polyphony carry undue baggage from over a century of musicological scholarship and that they should be approached afresh. The second chapter argued that there is no academic consensus on how musical reuse should be defined, both in the clausula-concept as well as beyond. The third chapter discussed effective methodological ways forward to the rigorous analysis of musical reuse in the ND repertory, by going beyond the diverse and diffuse definitions of "clausula" and the idiomatic criteria of melodic formulae. By the end of the third chapter a conceptual model had been created to this end, but the details and implementation were still to be ascertained.

The second part of this thesis progresses forward from this abstract thinking into more concrete areas of methodology, describing the processes, technologies, and systems that must be created to answer the research questions. Using the results of this, the following two chapters will be focused on implementation of the conceptual model developed in the first part of the thesis. It is in the second part that the digital humanities play a crucial role in enabling a rigorous analysis of musical reuse. However, before proceeding any further it is essential to understand how the digital humanities will be engaged with in this context, and what advantages it can bring to this study.

4.1. The digital humanities

The most common methodologies within the analysis of ND polyphony usually adhere to the principle of close reading, i.e. authors are typically engrossed in the analysis of a single setting of music, or a small cluster of interconnected pieces, focusing on uncovering relationships and interdependencies between them. This is an involved process — a journal article may revolve around a single set of examples — and this allows for a thorough exploration of the intricacies and contextual nuances within the selected pieces. Arguments made from the evidence created by a sustained analysis of a single set of pieces can be convincing, but the generalisability of these findings may be limited.[1] While the detailed analysis of a specific set of pieces can offer valuable contributions to the understanding of those compositions, applying the same arguments to a broader base can be challenging and potentially misleading, leading to what is essentially a sampling error. What may be true for a few select pieces does not necessarily hold for the repertory at large.[2]

To draw broader conclusions or make claims about the repertory as a whole, it is necessary to consider the repertory as a corpus, so-called "distant reading" (Moretti 2013). There are over 1500 settings of polyphony within the three main sources considered here, and to analyse each setting individually and its relation to every other setting through direct one-to-one comparisons is unfeasible (see "3.8. Methodological utility"). A corpus analysis is therefore required. The groundwork for such an analysis has been done through the many accomplished descriptions of the repertory — from the simple catalogues of Ludwig (1910) and van der Werf (1989) through to the painstaking comparative editions of Tischler (1988) — but an analysis of the repertory as an entirety does not exist.

To my knowledge, there are three existing studies that directly concern digital corpus analysis of ND polyphony: Erickson (1970), Falck (1990), and Flowers (2013).[3] These studies primarily examined rhythm and metrical stress in two-part music. The research completed as part of my Master's thesis (Stutter 2020) cannot reasonably be considered a digital corpus analysis as, although a digital corpus was created in the *ndp* dataset, the analysis itself was analogue and based on close reading. In Erickson's (1970) study he created a corpus of ND duplum using IBM punch cards. He then analysed statistical patterns in melody and rhythm and compared these with their relationship to harmony. He also extracted data on harmonic consonance and common melodic contours between organa (Erickson 1970, pp.62–83). Falck (1990) did not encode music in the computer directly, but more simply investigated harmony and stress in a sample of the incipits of some two-part conducti by counting the frequencies of intervals between voices, thereby determining likelihoods for individual stresses and consonances. Flowers (2013) expanded on this, analysing Falck's data using a more robust, Bayesian methodology and developing a model to predict stress and harmony. Additionally, Flowers utilised a Naïve Bayes classifier, trained on 210 so-called "synchronised" ordines in **F** (Flowers' terminology for passages where tenor and

[1]Exceptions to this are fortunately becoming more common nowadays. See, for example, the methodologies used in the *Josquin Research Project* (Rodin and Sapp 2010–) or the quantitative analysis performed on the *Ars nova* in Desmond (2018).

[2]A broader argument for the large-scale adoption of corpus studies to music analysis has been made by Neuwirth & Rohrmeier (2016). They argue that close reading alone is insufficient not only due to sampling errors but also methodological opacity and confirmation bias. This attitude has been more recently critiqued by Martin (2024), who emphasises that certain aspects of analysis are as yet unattainable by corpus study, even those supplemented by close reading.

[3]There was reportedly a study by Jane Moseley at Nottingham University (1986–1987) entitled "Source and Notation Studies, 1150–1200" using DARMS to match ligature patterns in ND polyphony (reported in Hewlett and Selfridge-Field 1987, p.113). However, it appears that no publications emerged from this project, and no physical records remain within the university. Huron also speaks of encoding ND polyphony using *Humdrum*, but this cannot have been more than a small sample (Shanahan 2020, p.12).

organal voice move together without overlapping) to predict rhythm in *cum littera* notation. The work of these previous studies is useful and important for providing groundwork for the possibility of analysing ND polyphony using computational methods, but the the focus of their studies — the interpretation of modal rhythm — bears little relevance to this study's focus on musical reuse. As described in "3.4.1. Graphical domain", aspects of rhythm are purposefully excluded from the dataset in this study, due to the difficulties of interpreting modal rhythm and the extent to which it can reliably be applied to ND polyphony. It is also important to note that although these studies provide interesting data from a sample of ND polyphony, their conclusions provide few insights beyond what can already be discerned from manual analysis. This study comprises a larger and broader analysis than these, in a fundamentally different aspect — set membership in musical reuse rather than rhythm, harmony, or melody — and begins directly at the graphical domain rather than a logical or analytical domain. The datasets created as part of these previous studies can therefore not be used as they begin from an already-editorialised logical domain rather than the graphical domain. A new dataset must be created.

A digital corpus analysis of a larger proportion of the repertory offers several advantages. Firstly, the ability to analyse everything all at once means that the analysis can capture a broader contextual understanding of the repertory, identifying overarching trends and connections between different settings that would otherwise be unfeasible or impossible by hand, and identifying areas of non-conformity that produce meaningful areas of study and analysis. Secondly, corpus analysis enables the identification of recurring patterns across the entire repertory and, with the appropriate parameters set, these patterns may point toward the identification of patterns of musical reuse. Thirdly, a distant reading creates data-driven hypotheses that are backed up by measurable evidence: it can be determined whether a pattern is significant or is just part of a general variation in the data.

Another often overlooked and underutilised advantage of harnessing the digital humanities is the ability to share the results of the research more readily in a manner that encourages exploration and engagement from both researchers and the wider community. Jensenius (2021) cited the FAIR principles (Wilkinson et al 2016) as best practice for sharing data in music research, and it is by these principles of Findable, Accessible, Interoperable, and Reusable that this project will continuously evaluate itself in the context of the digital humanities. However, it is important to note that the digital humanities community as a whole has not yet settled on how to accomplish this (Tóth-Czifra 2020).

A key part of the first two aspects of the FAIR principles will be in presenting the research in a way that is inherently more inviting and captivating to a broader audience than static prose on paper. Traditional research outputs (such as this one) often present information in a linear and static format, limiting the reader's ability to interact with and delve deeper into the subject matter. Digital tools, such as interactive websites, have the potential to provide a dynamic and immersive experience that can be more inviting and engaging to a user, as they make the analyses and editions more tangible when dynamically incorporated with other multimedia elements. Furthermore, the potential for accessibility and online availability of born-digital research makes it easily shareable, discoverable, and accessible to a wider audience. The convenience of accessing the research's data and results at any time and place enhances the likelihood of individuals encountering and engaging directly with the research. Such advantages are well known and well documented and need not be repeated here (Presner 2010).

4.2. Requirements

Balanced against these elevated ideas is the reality of the project as a small, single-author project with no funding. Often, projects create requirements by listing the Key User Requirements (KUR) of the system for each stakeholder. These requirements can be described as 'those that are absolutely mandatory' (Dick et al 2017, p.96). These requirements would then be supplemented by additional requirements until a full complement of prioritised requirements is generated for every interested party. Given the possibilities elicited by the digital humanities outlined in the previous section, it is quite easy for this requirement identification process to expand to an impractical size in this project. Several potentially interested stakeholders could be considered: researchers, performers, archivists and librarians, students and educational institutions, digital platforms, technology providers, as well as the wider public. A large, well-funded digital musicology project could and should be able to cater to a considerable proportion of these possible stakeholders, each with several key and additional requirements. However, it is important to state outright that this is not one of those projects. There are two steps that I therefore took to limit the size of the project: a) to prioritise the user requirements based on MoSCoW (Must, should, could, won't),[4] and b) to limit the stakeholders of this project to two key groups, one highly exclusive and one highly inclusive: me, and other possible interested parties respectively.

This second step may appear self-centred and imprecise, but framing the stakeholders this way in fact allowed the boundaries of the groups to be extremely well defined: there is the primary group of just-me, and a secondary group of everyone-apart-from-me. The first goal of this study is to learn more about the parameters of musical reuse within the ND repertory, and I am the (sole) researcher of the project. Any tools

[4]The "W" of MoSCoW is sometimes "wish" (see Dick et al 2017, p.98), sometimes "won't" (i.e. for requirements that explicitly fall outside of the project's scope) (see Achimugu et al 2014, p.581).

created for this project are created by me, to be used by me, for this research. To put any stakeholder apart from myself at the centre would be to stray far from the aims of this research: I intend the results of the research to be the primary research output, not the created tools themselves. If the project's digital tools do not comply at all with the FAIR principles for the secondary group (everyone else), it can still be said to be a successful project if the research questions can be answered. Any work done towards the requirements of a secondary group are a bonus: they are of course a net positive and should be aimed for, but should not be considered key to the success of the project. As a result, this project is also forgoing all end-user testing except for that done by myself. A broad base of user testing is something that would be vital to the success of a larger project, but there was no time or space to include it in this project. I will, however, in addition to Unit and System testing the technologies produced, be undertaking a degree of other end-user testing by evaluating the results through the persona of a technically proficient academic who wishes to explore some of the results of my work.

This practice of placing the self at the centre of the development process is colloquially known in the software development industry as "dogfooding" (Warren Harrison 2006), where software developers aim to use the very product that they are creating. This project has the further case of the key stages being an end to themselves: they only exist in order to fulfil the next stage of the process. This has the advantage of making the specification of requirements quite simple as there is often only one input and one output to the system or component. Each stage of the requirements moves the project closer to its goal.

What follows are the requirement statements that I created for both stakeholder groups considered. Firstly, and as a general over-arching aim, the system should enable the creation of analyses focused on musical reuse in ND polyphony, with a streamlined and user-friendly approach to data input and analysis. I have organised this requirements list by functional (F), non-functional (NF), and performance (P) requirements for both groups of stakeholders. For the primary stakeholder group (me), the system should meet the following requirements:

F1. Facsimile image import:
  a) Must provide the capability to import facsimile images of the manuscript sources for reference and visualisation purposes;
  b) Should support various image formats and sizes commonly used in the field of digital musicology.

F2. Manuscript tracing:
  a) Must provide a streamlined and intuitive interface to tracing elements from the facsimile images of the source manuscripts;
  b) Should enable users to tag the settings and orders of the systems and staves

within the traced manuscript images, while also providing the ability to add relevant metadata associated with each tagged element;

c) Could incorporate zooming, panning, and other image manipulation features to enhance the accuracy and ease of the tracing process.

F3. Musical reuse extraction:

a) Must develop algorithms or methods, including AI-based techniques, to assist in the extraction of musical reuse patterns from the traced manuscript images;

b) Should ensure that the extraction process maintains a high level of accuracy while minimising false positives and negatives;

c) Could provide capability for users to review, validate, and refine the extracted patterns, allowing for manual adjustments if needed.

F4. User authentication:

a) Must provide user authentication functionality, allowing users to create accounts and log in to the system;

b) Must implement appropriate security measures, such as secure password storage and transmission, to ensure the confidentiality and integrity of user credentials;

c) Must allow logged-in users to alter records within the database, including the individual attributes of models and relationships between sources, folios, facsimiles, systems, and staves;

d) Must enforce data validation and integrity checks to maintain the consistency and validity of records when alterations are made;

e) Should provide an intuitive and user-friendly interface for logged-in users to edit and update the model attributes and relationships, as well as alter the elements on staves through manual tracing;

f) Should provide appropriate error handling and validation feedback to users, notifying them of any issues or conflicts encountered during the alteration process;

g) Should support concurrent user access and handle potential conflicts or race conditions when multiple users attempt to alter the same records simultaneously;

h) Could employ access control mechanisms, such as role-based permissions, to restrict certain alteration capabilities to specific user roles or user groups.

NF1. User-friendly interface:

a) Should provide an intuitive and user-friendly interface that simplifies the manual tracing and extraction process;

b) Should provide clear instructions, guidance, and feedback to users during the tracing and extraction steps, ensuring ease of use and minimising errors;

c) Could use AI assistance as part of an OMR system to help automate the process.

NF2. Security and data privacy:

a) Must implement appropriate security measures to safeguard the integrity of the traced manuscript images and pattern analyses;

b) Must ensure compliance with relevant data protection regulations to protect user data.

P1. Efficient database browsing:

a) Must provide APIs or export functionality to allow seamless integration with external tools or scripts for programmatic access to the database;

b) Should respond quickly to user queries (under 100 milliseconds), and provide transcriptions and analyses in under 10 seconds.

P2. Scalability for multiple users:

a) Must be capable of supporting concurrent access from at least ten users;

b) Could incorporate user management features, such as access control, to support secure and controlled access for multiple users.

P3. Responsive user interface:

a) Must ensure that the system responds quickly to user input, providing a smooth and interactive user experience;

b) Should aim to achieve a response time of under 100 milliseconds for typical user interactions such as searching, browsing, and manipulating facsimile images or traced manuscript elements.

P4. Storage capability:

a) Must have sufficient storage capacity to accommodate the facsimile images of the central sources, estimated to be roughly 1GB in size;

b) Must be capable of storing the dataset associated with the traced manuscript images, estimated to be roughly 500MB in size;

c) Must support data backup and redundancy mechanisms to ensure data preservation and disaster recovery;

d) Should provide efficient data compression methods to minimise storage requirements without compromising data integrity or accuracy;

e) Won't include sources other than $\mathbf{W_1}$, $\mathbf{W_2}$, and **F**.

For the second stakeholder group (all other possible interested parties), the system should facilitate the exploration of the dataset as well as its analyses. The system should meet the following requirements:

F5. Analysis exploration:

a) Should enable users to navigate and explore the analyses conducted on the dataset;

b) Could provide tools and functionalities for filtering, sorting, and searching the analyses based on various criteria of the reuse parameters;

c) Won't provide a purely offline view of the data decoupled from the online web application and database.

F6. Comparative analysis:

a) Could allow users to compare and contrast different analyses or subsets of the dataset;

b) Could provide visualisations or summary statistics that highlight similarities, differences, and trends across multiple analyses.

F7. Interpretation support:

a) Could offer contextual information and explanatory resources that aid outside observers in interpreting the dataset and analyses effectively;

b) Could provide documentation, help guides, or tooltips that clarify the terminology, methodologies, and assumptions used in the analyses.

F8. Collaboration and sharing:

a) Could allow for the export of visualisations, reports, or selected subsets of the dataset in formats suitable for further analysis or presentation;

b) Could enable users to share specific analyses or visualisations with others for collaboration or scholarly discourse.

NF3. Dataset access:

a) Should implement appropriate access controls and permissions to ensure data security and integrity;

b) Should provide a user-friendly interface that allows access to the dataset containing traced manuscript images and associated metadata.

NF4. Visualisation and exploration:

a) Could generate visual representations of the musical reuse patterns, such as notations or graphical renditions;

b) Could implement interactive features that enable users to zoom, scroll, and manipulate the visual representations to gain different perspectives on the data.

NF5. User-friendly interface:

a) Should provide an intuitive interface that requires minimal training or technical expertise to navigate and explore the dataset and analyses;

b) Should ensure that the interfaces incorporates accessible contextual guidance and clear instructions, aiding users in seamlessly navigating and utilising the system to explore the dataset and analyses effectively.

NF6. Open-source software:

a) Should enhance transparency and reproducibility, allowing researchers to validate and verify the software's algorithms, methodologies, and results;
b) Could facilitate the adoption and extension of the software in related research projects, avoiding proprietary software dependencies and promoting its widespread use.

NF7. Ease of maintainability and future-proofing:
a) Should be designed with clean and modular code structure, adhering to established coding standards and best practices, to facilitate ease of maintainability;
b) Should include comprehensive documentation, including code documentation, system architecture, and configuration details, to support efficient maintenance, troubleshooting, and knowledge transfer;
c) Should employ version control systems to track and manage changes, providing a historical record of modifications and enabling efficient collaboration among maintainers;
d) Should be designed with flexibility and extensibility in mind, allowing for easy integration of new functionalities, technologies, or components as future needs arise;
e) Should prioritise compatibility with emerging technologies and industry standards;
f) Should consider the ongoing availability of necessary resources, such as software dependencies, libraries, or APIs, to ensure continued support and maintenance;
g) Could employ scalable and modular architecture, allowing for efficient scaling and future expansion to accommodate increasing user demands or evolving requirements;
h) Could incorporate automated testing and deployment mechanisms to ensure that changes and updates can be implemented with minimal risk and disruption.

P5. Scalability for more users:
a) Should monitor system performance and resource utilisation to proactively identify bottlenecks or limitations, allowing for proactive scaling and optimisation to maintain a seamless user experience under varying user loads;
b) Could be capable of handling increased user load and scaling up to accommodate hundreds of concurrent users;
c) Could utilise load balancing techniques and distributed processing capabilities to distribute user requests efficiently across multiple servers or

instances;

d) Could incorporate horizontal scalability by adding additional resources or nodes to the system's infrastructure to ensure optimal performance and response times, even with high user load.

4.3. Choosing technology

A sizeable proportion of the following two chapters will be heavily involved with the adoption and description of technologies to implement a system that fulfils the above requirements. As the focus of this thesis is on process rather than technological infrastructure, the choice of technologies to implement a system that fulfils the above requirements — and the rationale for that choice — will not be discussed in detail. There are as many ways of creating systems as there are systems: very rarely is it that there is only a single technology capable of completing the task required. Once requirement lists were drawn up for each component of the system and technology evaluated, it was most often found that several technologies would be more than suitable for the task. This is overall a good thing, as it affords greater flexibility and choice in the selection of technology. There were three main principles that guided my decisions, and I will clarify them here.

The first principle is leveraging existing expertise. I did not want to spend a substantial portion of my time during the research phase learning new technologies that replicated already-existing knowledge. Granted, new technologies had to be learned in areas that I had little experience in, but I had the advantage of coming to this research already with a broad range of skills within the domain. It would have made little sense to disregard these skills for the sake of novelty or at best a minor advantage.

The second principle is to avoid premature optimisation. It is often tempting to select technologies simply for their claims of scalability and performance. However, such technologies are often more difficult to use or do not come with more complex or higher-level features as standard. For example, it is perfectly possible to write a web application in lower-level languages such as C or C++, containerised within a replicated swarm of load-balanced instances. This would be extremely performant and capable of scaling to many thousands of users. However, it takes many times longer to write and debug such programs, and to administer the architecture around them. It is not envisaged that this particular implementation will be scaled out, and so the number of users is small. As such, any scaling requirements will be left to a future project if ever needed. Regardless, any performance savings made by writing the necessary code in lower-level languages would be completely eclipsed by the true data bottleneck of the application: that of querying databases.

The final principle is that of maintainability, support, and future-proofing. This is a crucial principle as it emphasises the importance of choosing technologies that can be easily maintained and supported over the long term. There may be technologies that suit the purposes exactly, but are unmaintained with little documentation, and/or without a community to help keep the project running. A useful metric I used here was to use technology that has been continuously maintained over the last ten years. Maintainability minimises technical debt and reduces the likelihood of encountering compatibility issues or dependencies on outdated or obsolete components.

An influential essay in current software development practice is the mantra "Choose boring technology" (McKinley 2015). The basic premise is that if you are trying to do something radically different, you should not also hobble yourself by trying to do that using unproven technology. Rather than equivocating on "the best tool for the job", technologies should be picked that allow for solving a broad base of problems with minimal fuss. If you are trying to solve an already solved problem in a better or more elegant way, then exciting new technology should be encouraged. However, typical problems are new in terms of business rather than in technology, and so the technology should not also be new. A related mantra is "You are not Google" (Nova 2017). While tech giants such as Google may develop innovative technologies to address their specific global scaling challenges, it is important to recognise that such specialised solutions are rarely applicable to the needs of most projects. It is highly unlikely that any of these tailored technologies will be needed. Leveraging standard tools and frameworks is often more than sufficient to meet project requirements effectively and efficiently. These are the overriding themes that I have gone by here. The research questions address unproven methodologies of researching ND polyphony using digital technologies. I therefore should be using simple and proven technology where possible.

4.4. The Clausula Archive of the Notre Dame Repertory

The system created from the requirements is the *Clausula Archive of the Notre Dame Repertory* (CANDR).[5] CANDR is a system consisting of two main parts: a web application for the browsing, editing, import and export of ND polyphony as facsimile images and symbolic notation, accompanied by a programming toolkit that interfaces with the web application to create computer-aided analyses of musical reuse within the repertory. CANDR provides a comprehensive platform designed for scholarly investigation and exploration of ND polyphony, catering to the needs of researchers and musicologists by providing a diverse array of useful tools and resources. CANDR therefore represents a significant advancement in the field of medieval digital

[5] CANDR can currently be accessed at <https://www.candr.org.uk/>.

High-level view of the architecture of CANDR

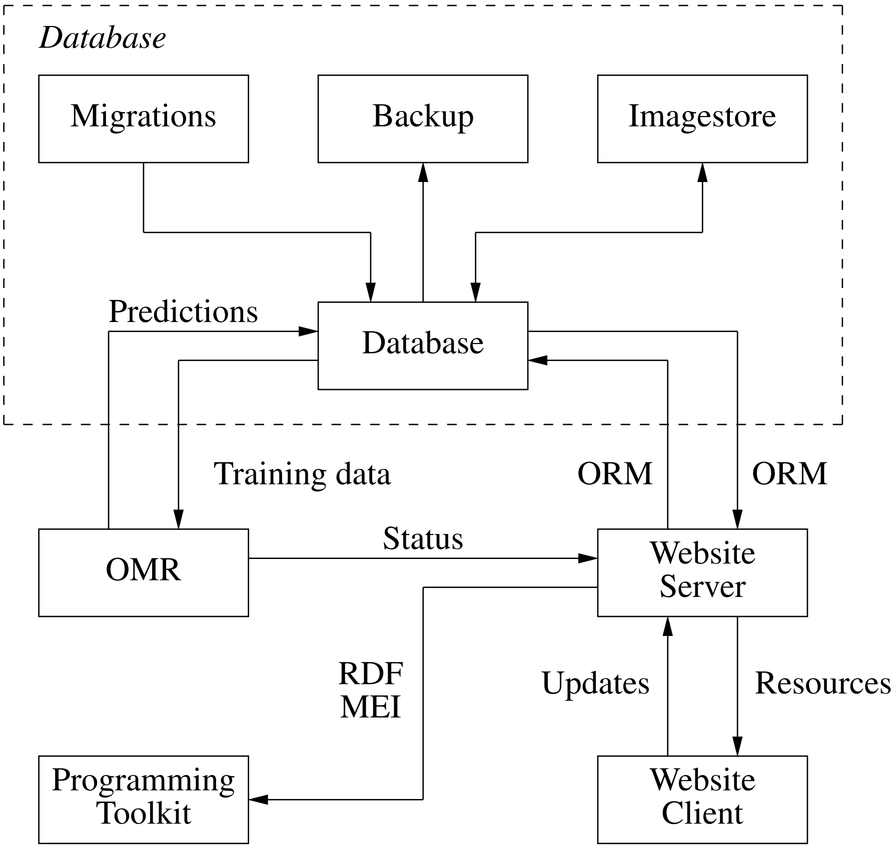

Fig. 4.1

musicology.

The web application serves as a centralised hub that facilitates browsing, editing, and management of the extensive collection of ND polyphony. Through a user-friendly interface, scholars can seamlessly navigate through the repertory, delving into the minutiae of ND polyphony in a hierarchical framework that facilitates an in-depth examination at every level. The application enables users to begin at an overview of manuscripts and gradually refine their focus to scrutinise specific settings and even individual notes. By presenting a broad swathe of the repertory, this multi-layered approach allows for a thorough exploration of the more intricate facets of the music, empowering researchers to gain a nuanced understanding of the polyphonic tradition.

The programming toolkit, tightly integrated with the web application through an open API (Application Programming Interface), equips researchers with the necessary tools to conduct advanced computer-aided analyses of musical reuse within the ND repertory. Leveraging completely novel and state-of-the-art algorithms and data processing techniques in the field, the toolkit allows for the identification and comparison of patterns of musical reuse, elucidating intricate patterns of polyphonic transformations across multiple settings. This analytical capability opens new avenues

Requirement–component matrix

	Database	**Website server**	**Website client**	**OMR**	**Toolkit**
F1	×	×			
F2			×		
F3					×
F4		×	×		
NF1			×	×	
NF2	×	×			
P1	×	×			
P2		×			
P3			×		
P4	×				
F5			×		
F6					×
F7					×
F8		×			×
NF3	×	×			
NF4					×
NF5			×		
NF6	×	×	×	×	×
NF7	×	×	×	×	×
P5	×	×			

Tbl. 4.1

for understanding the evolution, influences, and stylistic nuances within ND polyphony.

Furthermore, CANDR embraces the principles of open science and collaboration. The system's architecture and data formats are designed to be as open and interoperable as posible, enabling effortless integration with other research tools and facilitating the sharing of datasets and analyses. By adopting open standards, CANDR encourages a vibrant scholarly community and fosters collaborative research initiatives.

Fig. 4.1 depicts a high-level representation of the architecture of CANDR, consisting of five major components: database, OMR (Optical Music Recognition), website server, website client, and the programming toolkit. The database consists of four sub-components: the database server itself, a separate image store for storing facsimile images and thumbnails, a regular backup solution to maintain data integrity, and a set of migrations to describe the database's schema. The database is tightly connected to the website server using an Object Relational Mapper (ORM) which describes the schema in object-oriented code: the objects that exist in the database, their attributes and types, and defines some common tasks that can be performed on them and their relationships. The website server provides an HTTP frontend to the database which serves as a CRUD (Create, Read, Update, Delete) model for the database. Many of the client's functionalities are simple calls to alter database fields. The client has more advanced functionality, however, in a *JavaScript* web application graphical user interface (GUI) that serves as the main browsing and editing interface. This application stores its state locally, and regularly pushes updates back to the server and database. This is common functionality for a web application and therefore many of its features will not be described in depth. Tbl. 4.1 presents the way in which individual components have been delegated to fulfil each of the requirements.

The database and website server are also linked through a novel OMR application designed to accelerate data keyboarding by training a convolutional neural network (CNN) on the current database contents. It continuously cycles through staves of music in the database and updates its model based on ground truth, emitting predictions for each un-transcribed stave. While the model is not perfect, it provides valuable predictions that serve as a solid starting point for transcription and significantly reduces the manual effort required for transcription. The website client has the functionality to copy the OMR predictions onto the currently edited stave, update and correct the data, and save the result as new ground truth for the OMR application to update its model.

These compose the "online" component of CANDR. The "offline" component comprises the programming toolkit which leverages several completely novel methodologies for conducting corpus analysis on music, based on the data structures and methodologies reviewed in "Chapter 3: Data models". The toolkit has the capability to access the website's API and retrieve both RDF (Resource Description Framework) metadata as well as symbolic transcriptions of passages of music in MEI. It then provides a number of general and parameterised programming modules to create analyses of musical reuse, along with a data framework for organising, structuring, executing, and saving the results.

4.5. Ontology

To allow users to drill down from the very widest concept covered in the system (the "source") down to the very smallest (the "item"), CANDR employs a custom hierarchical ontology for its bibliographic domain (here called the first-level schema) and a separate ontology for the graphical (the second-level schema). The first-level schema contains the entities: source, folio, facsimile, system, and stave (see Fig. 4.2). These are generally organised in many-to-one relationships. For example, a source contains many folios, a folio many facsimiles (but typically only one), facsimiles many

CANDR's first-level schema

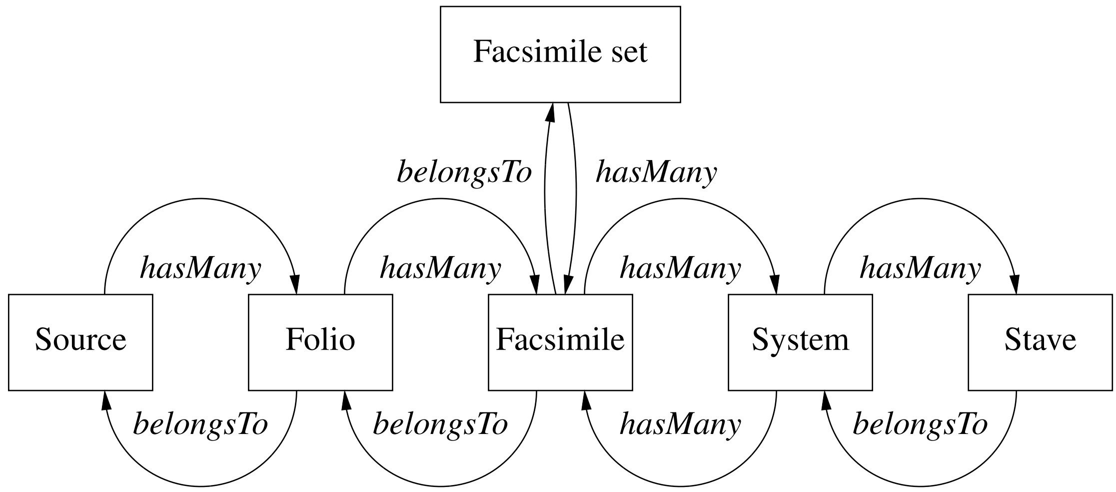

Fig. 4.2

systems, and systems many staves. This allows CANDR to be able to assign unique identifiers to each stave based on its system. We can know, for example, that the stave with ID number 75 is the third stave on the first system of the only facsimile of folio 5v of $\mathbf{W}_2$, and that it can be found at the path `/browse/stave/75/`.

This general relationship is broken for two exceptions: the facsimile set, and the system–facsimile relationship. In the first case, facsimiles belong not only to folios, but facsimiles can also belong to facsimile sets (facsimiles taken from a single source and grouped together in the system). A facsimile set has its own name, an upstream URI where the images were sourced from, as well as copyright and licencing information that must be adhered to. Linking all facsimiles gained from a single source into a facsimile set allows for licencing that does not allow for further distribution to be marked as such, and the images hidden from view for unauthenticated users. Secondly, there are rare cases where systems appear on multiple facsimiles: this can be the case if there are multiple facsimile sets: wide as well as detailed shots, photographs of the manuscript surface from different angles, or where systems of one original folio may have been split into multiple fragments. Not only, then, does a facsimile have many systems, but a system can belong to many facsimiles. This makes transcription difficult, so in most cases a single "transcription" facsimile is indicated as the image from which the system will be traced from.

The second-level schema contains the entities: accidental, clef, divisione, ligature, note, stave line, syllable, as well as editorial counterparts (see Fig. 4.3). These represent the graphical domain, and so are arranged exactly as outlined in "3.4.1. Graphical domain". They all fall under the umbrella of a generic "item", such that any element on a stave can be queried by a unique ID number. The exception to this is the ligature, CANDR's second-level schema

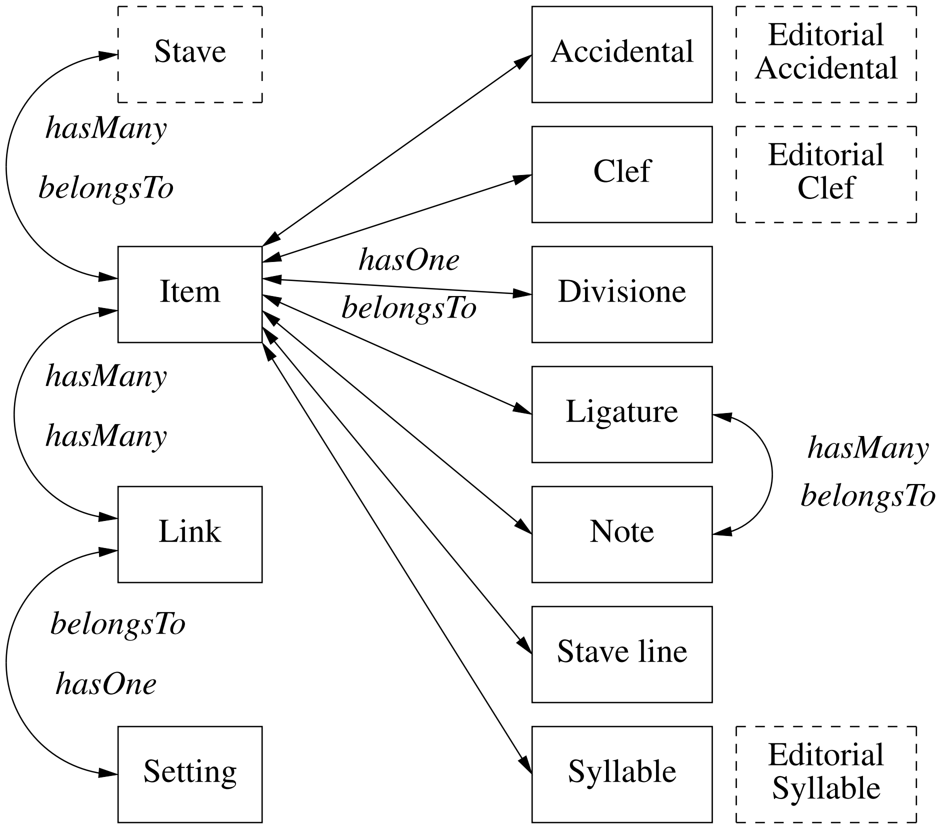

Fig. 4.3

which can also contain many notes. The item entity links back to the first-level schema by appearing as an unordered set within a stave. Items can additionally be ordered by the *x*-coordinate of their position, reading left to right, but this is not strictly required. The editorial synchronisation between staves takes place through a bespoke "link" record, which can contain many items — although each must be on a different stave as it makes little sense to link two items on one stave — and a link indicates that those items are deemed to reasonably occur simultaneously. This link entity is virtual: it does not relate to any item on the page, but is instead an editorial parameter inferred by hand from visual clues such as position, cadence, harmony, and vertical alignment in the manuscript. The bibliographical domain enters through the concept of the "setting". A setting here is simply a named tag added to a link, indicating that a setting of music begins at this link. A transcribing user, having linked the opening of a setting between its constituent staves within a system, can then tag that link as a named setting (typically its text incipit). It is then up to a logical domain process to symbolically interpret all the music between two setting tags in left-to-right order as a singular "setting".

4.6. An overview of CANDR

A typical first page a user would navigate to using the web frontend of CANDR would be the "Browse sources" page (Fig. 4.4). From here, a user can view the names of the sources that CANDR catalogues. Clicking on a source itself displays a list of the folios within that source (Fig. 4.5). Then clicking on a folio displays a list of facsimiles associated with that folio (Fig. 4.6). Then clicking on a facsimile displays the facsimile itself in full resolution, overlaid with any defined systems on that system as free quadrilaterals (Fig. 4.7). Then clicking on a system quadrilateral displays a straightened version of that system with any defined staves overlaid once again as free quadrilaterals (Fig. 4.8). Finally, clicking on a stave brings up a straightened version of that stave overlaid with any second-level schema elements such as notes or divisiones. (Fig. 4.9). The visibility of the elements can be toggled by class. In this way, a user can descend the hierarchy and browse the sources at any level of granularity from the source level down to the individual item level. The reverse can be performed by starting at any level and viewing the breadcrumb navigation at the bottom of each page to retrace the steps back to the top level.

Authenticated users are granted extended functionality by the ability to alter the database. By selecting the "Login" option in the top menu, the user is prompted for a username and password (Fig. 4.10). Once logged in, any items connected to a model attribute or relationship also have a pencil icon and delete icon alongside (Fig. 4.11). Any spaces for further records (such as more folios to be added to a source) are indicated by a plus icon. Clicking an edit icon brings the user to a page for editing that record, a delete icon prompts the user to confirm deletion of the record, and the plus icon adds a new record to the database and takes the user to the page for editing that new record.

The facsimile and system editor pages extend the functionality of the browsing pages by adding draggable handles to the corners of the system and stave quadrilaterals (Fig. 4.12). This allows a logged-in user to alter the bounds of these items, as well as their ordering on the page. Additionally, the stave editor is augmented with extra functions for adding, editing, and removing elements from the stave, or copying the predictions of the OMR application onto the stave (Fig. 4.13). Clicking and dragging the handles of stave elements moves and shapes them on the image. Selecting each element brings up a list below of all its attributes in the database, such as position and type. The editor constantly sends live updates back to the server as the stave is edited.

Finally, this transcription process is streamlined by a companion "transcribing wizard" module that keeps track of the user's position in the database, and prompts the user to perform the next logical action, such as transcribing the next stave. Once all the staves in a system are marked as transcribed, the editor is then prompted to use a separate interface to link items from parallel staves in a system together as synchronous (Fig. 4.14). From here, an additional “Add setting” button allows a link to be transformed into a named setting tag.

Most items from the first-level schema also have a corresponding `/mei` endpoint that generates real-time and fully automated logical domain transcriptions of the tracings into MEI. As soon as a stave is updated, a message is propagated through the system to update all transcriptions that include this stave. This way, transcriptions are guaranteed to not be stale, and MEI exported from the system always reflects the latest data.

The programming toolkit uses these endpoints to download transcriptions from CANDR. The following example *Python* scripts are designed to be run sequentially to fetch, parse, and analyse the dataset. The first script uses the programming toolkit to download all settings from CANDR and saves the results in a file for further processing:

```
#!/usr/bin/env python3
from candr_analysis.scrapers.candr.rdf import setting
from candr_analysis.transcribers.candr import MEI
from candr_analysis.utilities import pipeline, progress
from candr_analysis.transformers import pickle

prog = progress.CLIProgress()
processes = pipeline.Pipeline(progress=prog)

processes.add(
    setting.Setting(name="RDFSetting", progress=prog),
    MEI.Setting(name="MEITranscriber", progress=prog),
    pickle.PickleSaver("settings.pkl", flush=True)
)
processes.execute()
```

This script takes a few minutes to execute, depending on the speed of the connection to the CANDR server and which settings need re-calculating. The next script parses the MEI files into notation graphs for analysis:

```
#!/usr/bin/env python3
from multiprocessing import cpu_count
from candr_analysis.parsers import MEI
from candr_analysis.utilities import pipeline, progress
from candr_analysis.transformers import pickle

prog = progress.CLIProgress()
processes = pipeline.Pipeline(progress=prog)

processes.add(
    pickle.PickleLoader("settings.pkl", progress=prog),
    pipeline.Multiplexer(
        [
            MEI.MEIParser(
```

```
                editorial=False,
                accidentals=True,
                ligatures=False
            )
            for _ in range(cpu_count())
        ]
    ),
    pickle.PickleSaver("parsed.pkl", flush=True)
)
processes.execute()
```

This script takes roughly half an hour to execute. The next script takes the parsed MEI notation graphs and creates ego graphs of size 5. For every node in each graph, all the nodes in that node's neighbourhood are extracted as a subgraph. These are then postprocessed into formats suitable for the particular machine learning libraries used in the next stage, and unique features are stripped from each node (such as each node's unique ID and location on the stave) so that nodes can be more easily compared:

```
#!/usr/bin/env python3
from functools import partial
from candr_analysis.utilities import pipeline, progress
from candr_analysis.transformers import pickle
from candr_analysis.transformers.MEI import apply_func, ego_graph

prog = progress.CLIProgress()
processes = pipeline.Pipeline(progress=prog)

processes.add(
    pickle.PickleLoader("parsed.pkl", progress=prog),
    ego_graph.EgoGraph(5),
    ego_graph.EgoGraphsToData(
        partial(
            apply_func.replace_tuple_attributes,
            strip_attrs=('uniqid', 'ligature_type', 'editorial', 'len', 'loc'),
            replace=None
        )
    ),
    pickle.PickleSaver("ego_graphs.pkl", flush=True)
)
processes.execute()
```

This script generates roughly 400,000 ego graphs in the correct format for further processing, and takes roughly an hour to execute. The next script uses this data as training and test data for a graph embedding methodology, generating embedding vectors of size 16 for each node using a node prediction task trained over 10 epochs:

```
#!/usr/bin/env python3
from candr_analysis.utilities import pipeline, progress
from candr_analysis.analysis.graph import graph_network, graph_statistics
```

CANDR's "Browse source" entrypoint

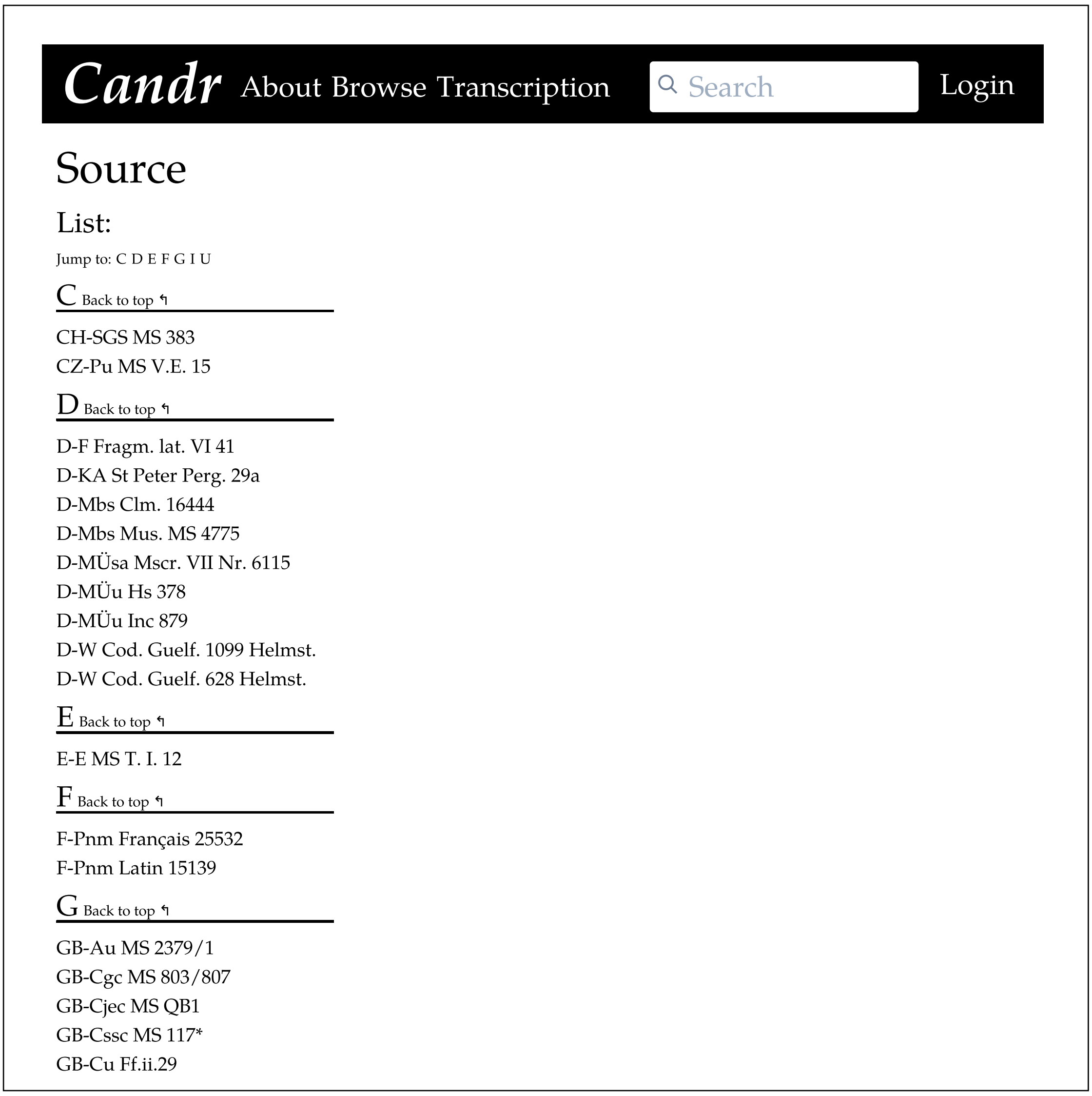

Fig. 4.4

```
prog = progress.CLIProgress()
processes = pipeline.Pipeline(progress=prog)

processes.add(
    pickle.PickleLoader("ego_graphs.pkl", progress=prog),
    graph_network.GraphRGCNNodePred(
        embedding_size=16,
        batch_size=40,
        dropout=0.2,
        learning_rate=0.001,
        epochs=10
    ),
    graph_statistics.GraphStatisticSaver(folder='log_folder/')
)
processes.execute()
```

As part of this process, the script generates embeddings for each node, so that a node's

Browsing $\mathbf{W_2}$

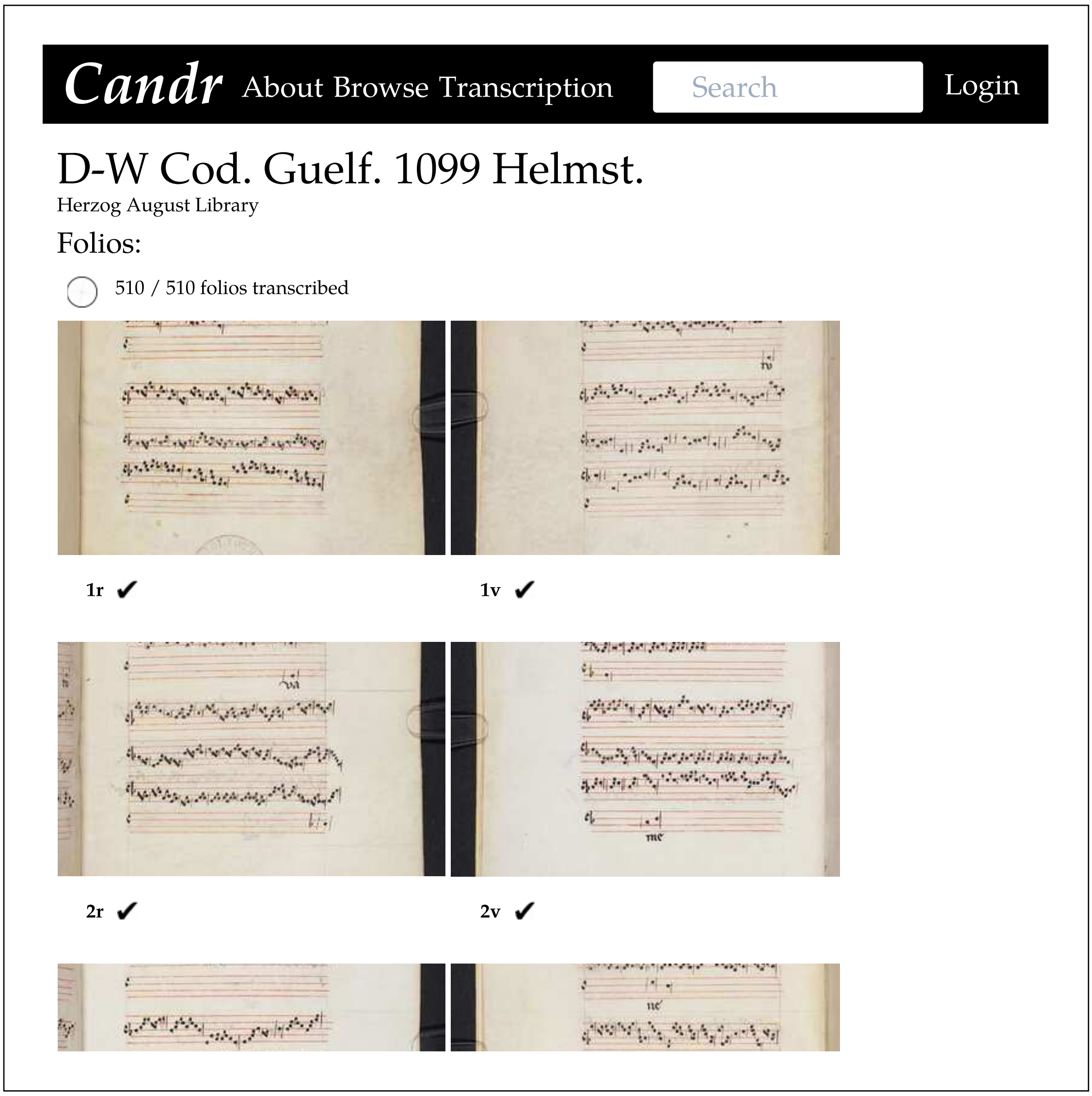

Fig. 4.5

embedding can be instantly looked up in a Comma Separated Values (CSV) file without having to be recomputed. A final script, existing outside of this pipeline ecosystem, then uses both the parsed notation graphs and generated embeddings to generate best-fit alignments of every setting to every other setting. From this dataset, plots can be created of similarity within each pair of settings. A Hidden Markov Model (HMM) can then be trained to discern between pairs that have no similarities and pairs that contain passages of high similarity. From this methodology, the results of the study presented in "Chapter 7: Results and Analysis" are created. The exact methodology will be described in more detail in "Chapter 6: Offline".

This chapter has made the case for applying digital humanities methodologies to the research questions of this study, primarily the issues of scale, transparency, and reproducibility. Through the initial research conducted in the first part of this thesis, a

Browsing $\mathbf{W}_2$ f.5v

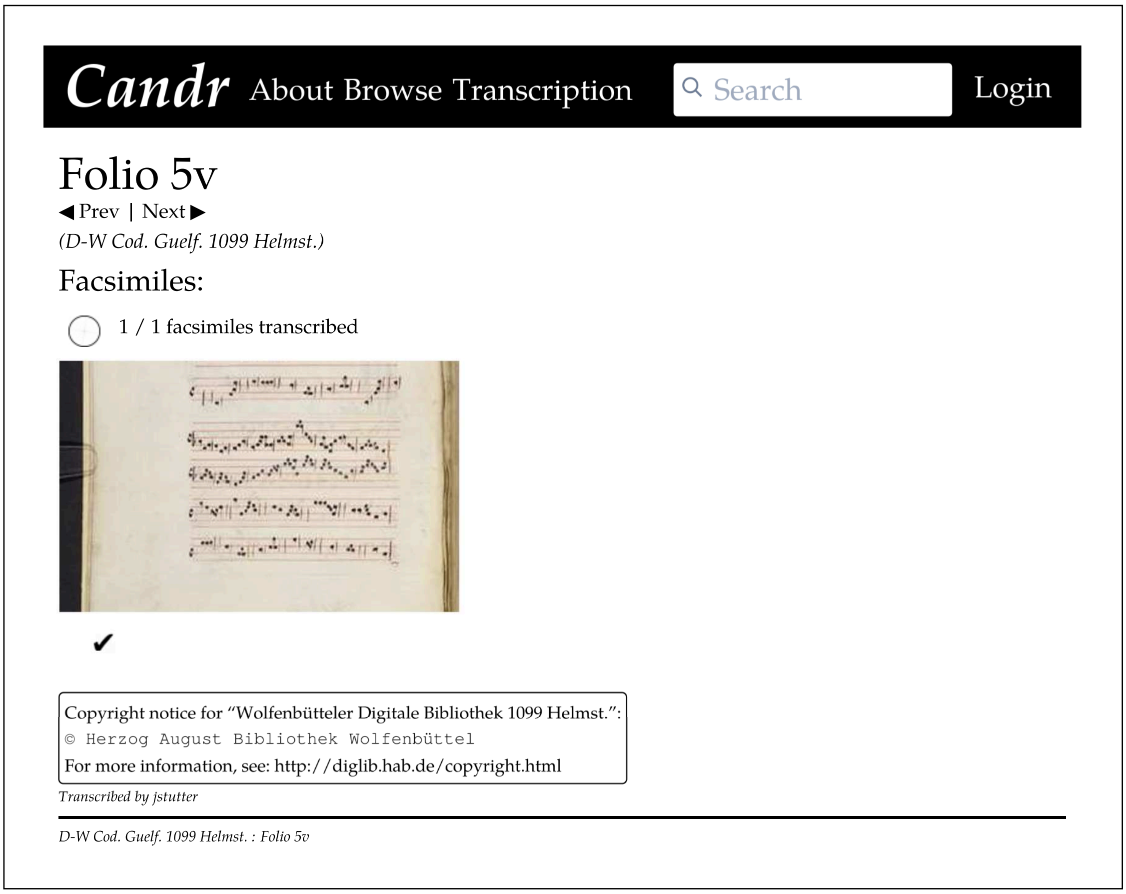

Fig. 4.6

set of requirements for a digital musicology system capable of addressing the “musical reuse” research question has been compiled. Subsequently, the guiding principles for selecting appropriate technology for the system have been outlined. Finally, the CANDR system has been introduced as a solution that satisfies the identified requirements, accompanied by a brief overview of its functionality. The following two chapters will delve into comprehensive explanations of the methodologies employed to implement both the online and offline parts of the system.

Browsing $\mathbf{W_2}$ f.5v facsimile

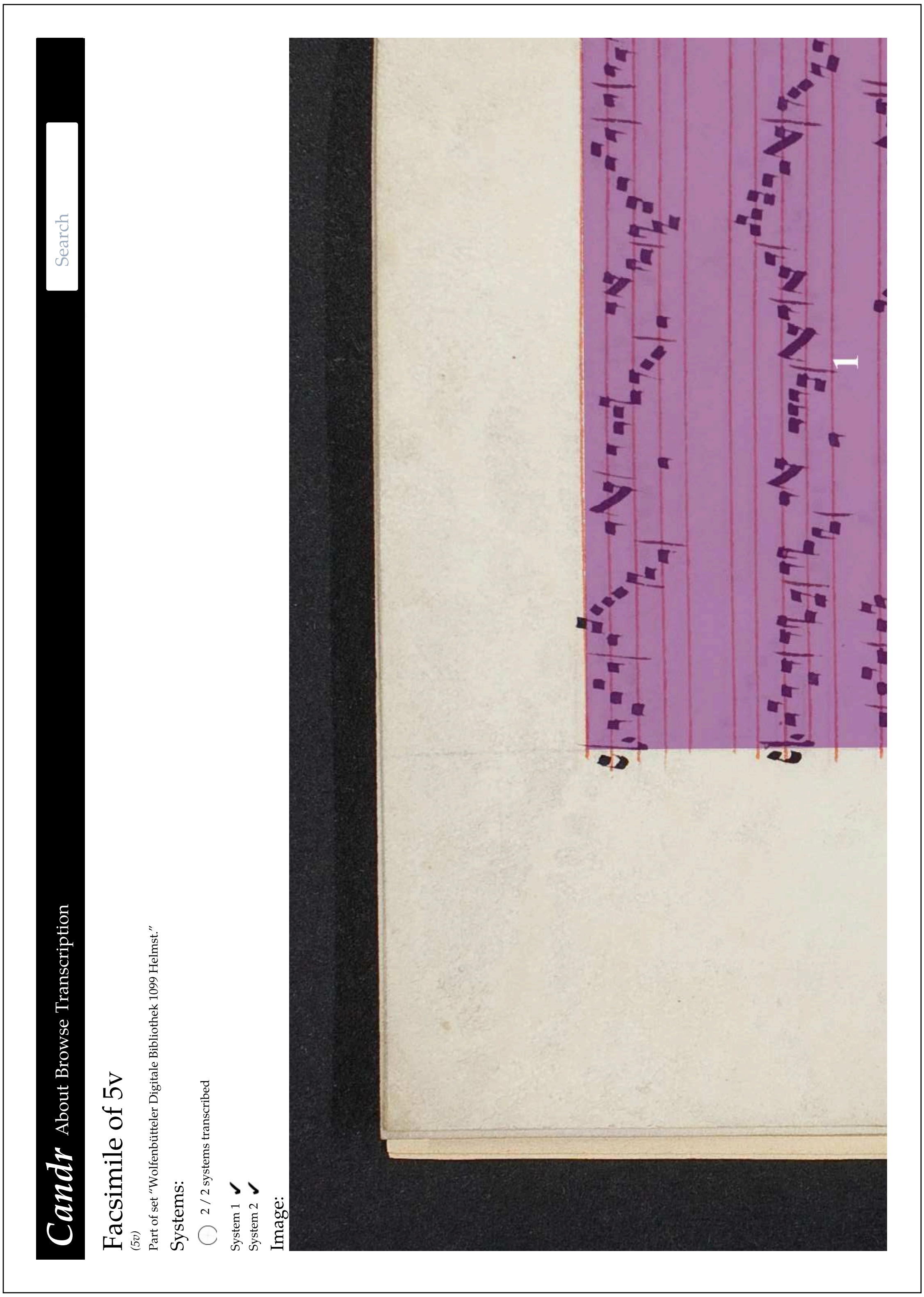

Fig. 4.7

Browsing system of **W**$_2$ f.5v

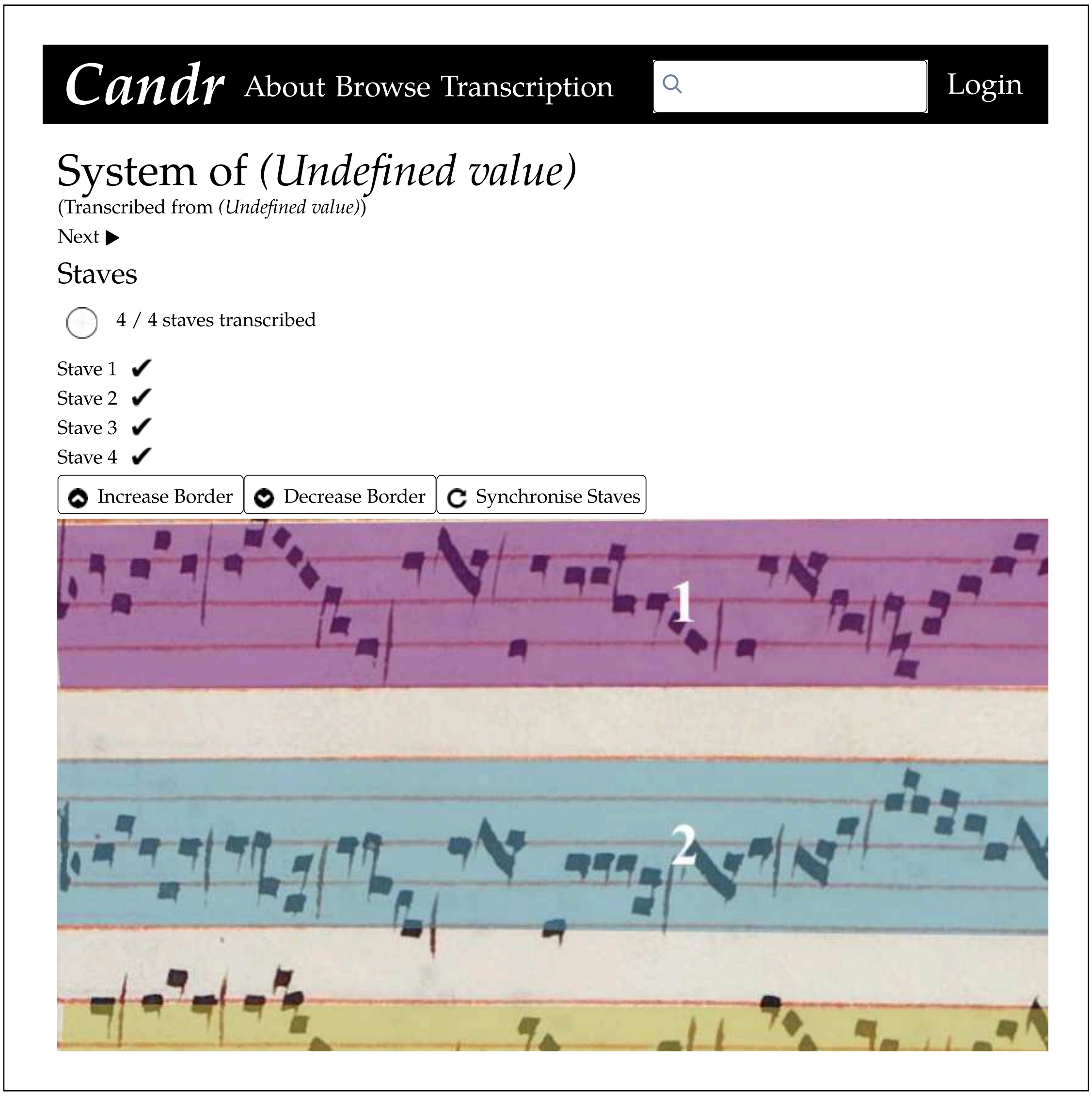

Fig. 4.8

Browsing stave of $\mathbf{W}_2$ f.5v (ID #75)

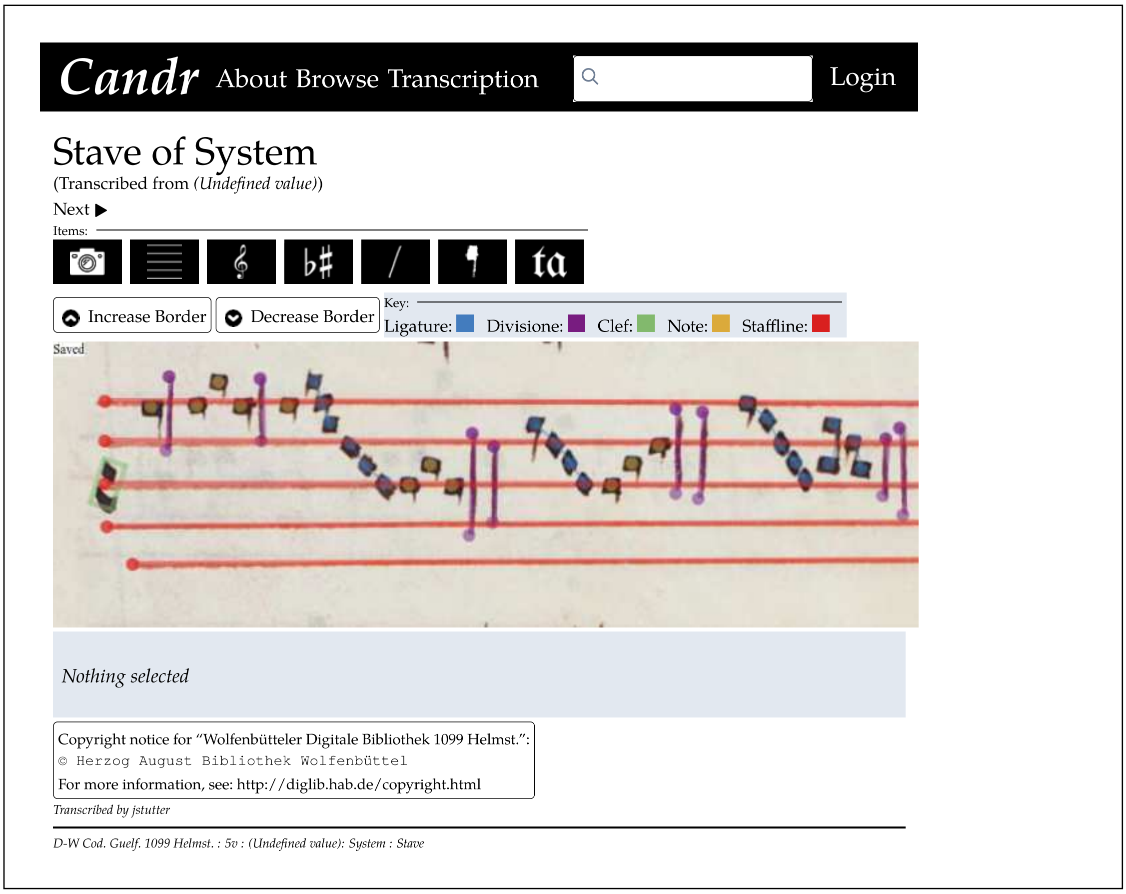

Fig. 4.9

CANDR login prompt

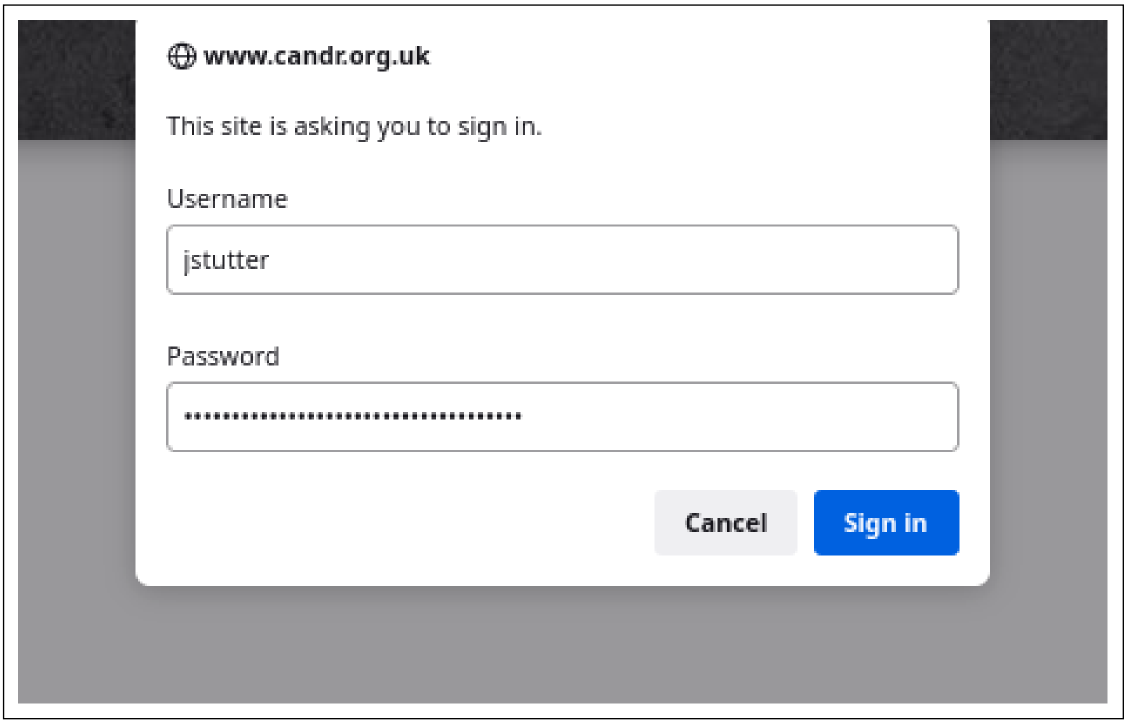

Fig. 4.10

Editing source view with editing icons

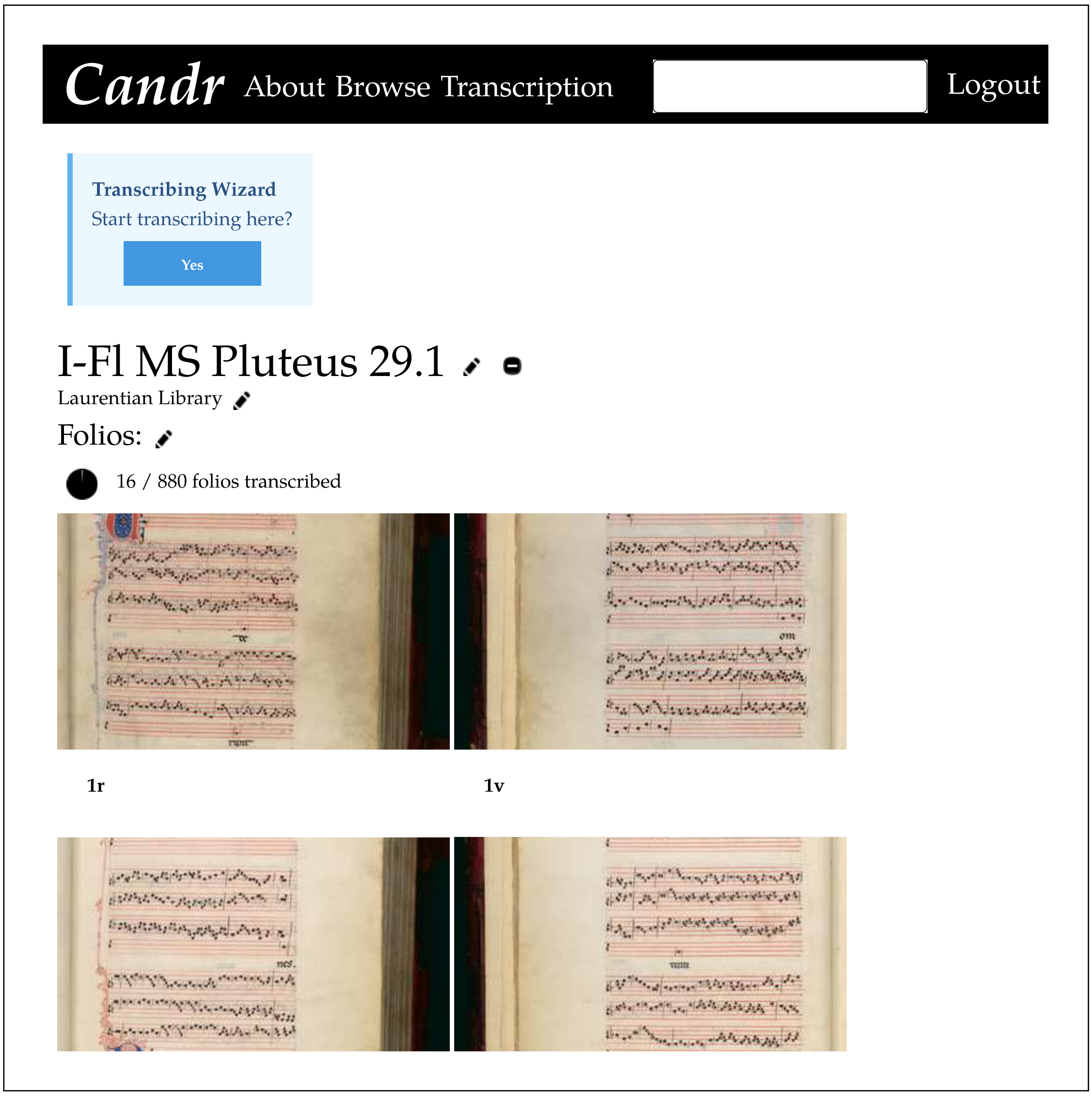

Fig. 4.11

Editing facsimile systems

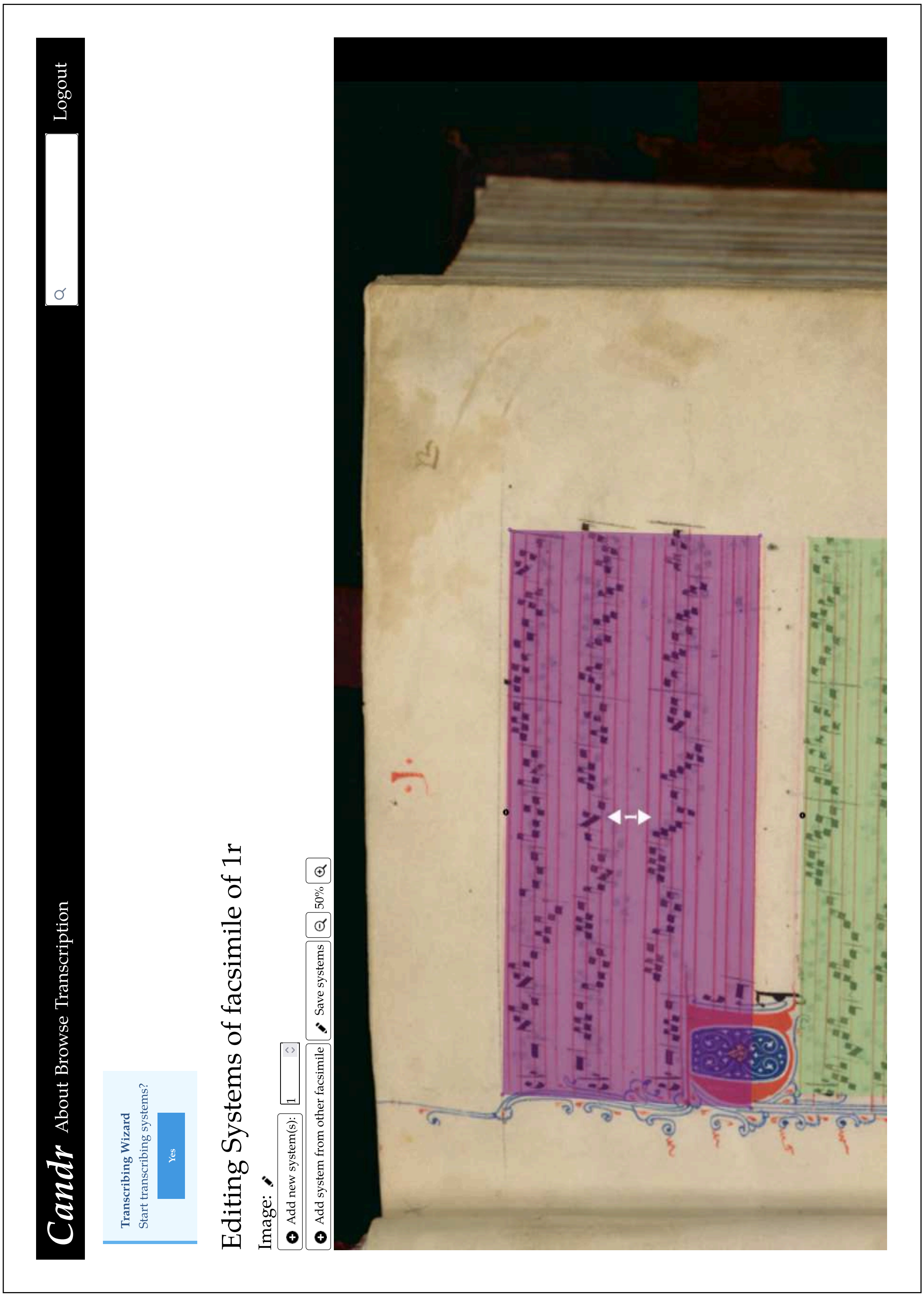

Fig. 4.12

Editing a stave

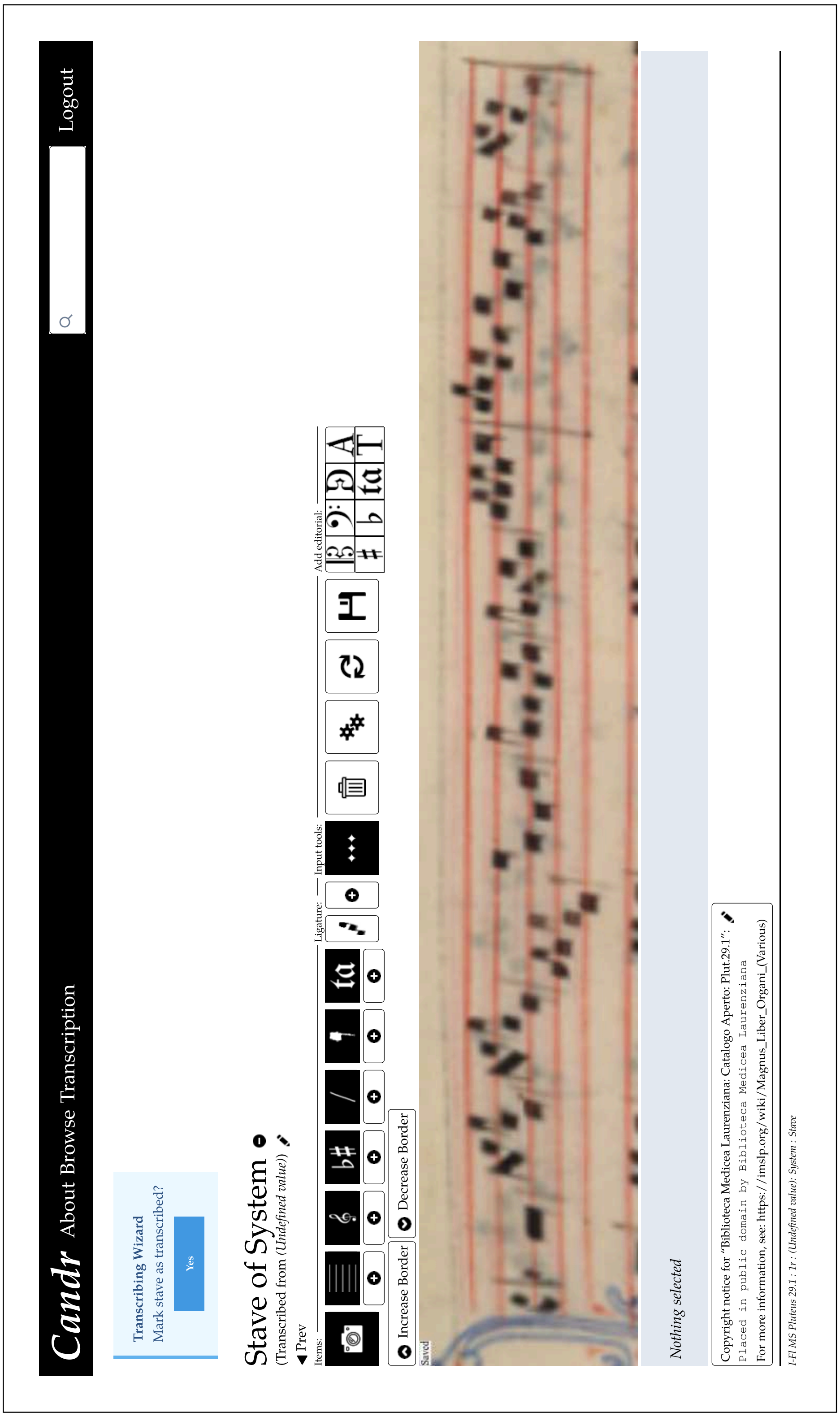

Fig. 4.13

Editing system synchronisation

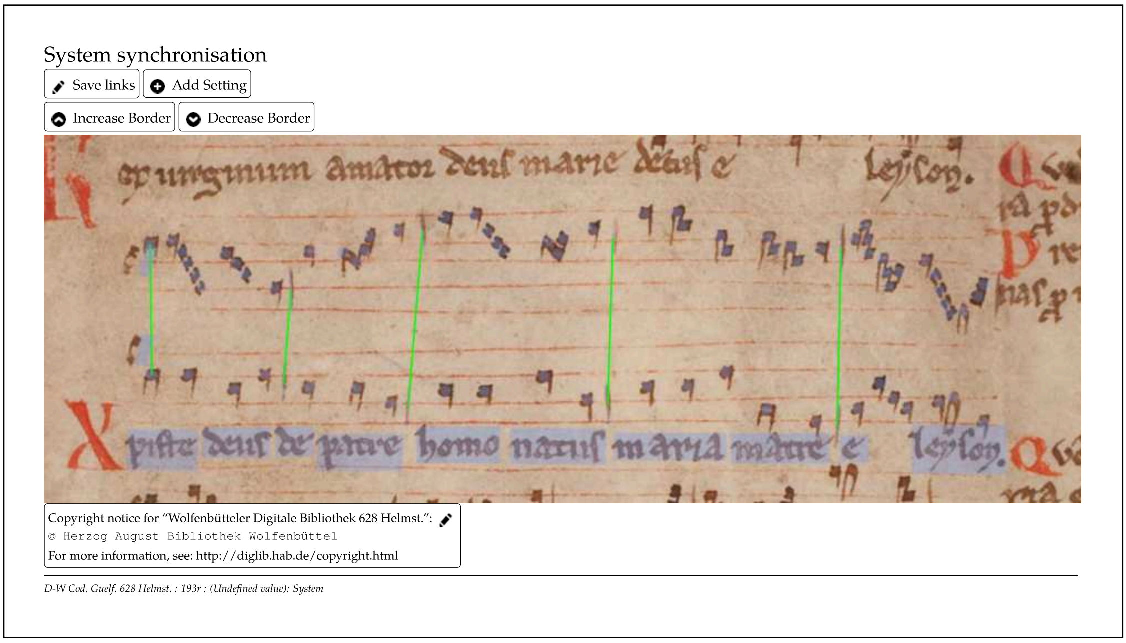

Fig. 4.14

Chapter 5
Online

Having identified the project's needs, and in the absence of existing tools that fully support those needs, an appropriate series of new tools were designed and built. In order to support the data entry requirements for this research, this toolset included a database, website server, website client, and OMR system. The purposes of these components are twofold: a) to enable and streamline the input process for the graphical domain dataset, and b) to present an intuitive interface for viewing and editing the dataset.

Although all of the components are vital to the operation of the system, it is important to acknowledge that certain aspects of the system are more commonplace or conventional in nature. Therefore, this exposition will predominantly emphasise only the innovative approaches, offering comprehensive explanations of their implementation, while only briefly touching upon customary elements that align with standard practices within similar projects. For example, the OMR application was designed from scratch for this purpose and so will be explained in detail, situating it within recent research on the use of neural networks for this task. On the other hand, the use of an RDBMS (Relational Database Management System) to store the dataset is arguably the most conventional option available and its usage here is not extraordinary, therefore its functionality will not be explained save what is necessary to reproduce the research.

For the sake of clarity, each component section will take a similar structure. There will first be a description of the component, including its required functionality and inputs/outputs. There is then an analysis of some currently available technologies that could implement these features, followed by the choice made, along with a short description outlining the reasons for that choice. Finally, I will explain how that technology is implemented in the system, as well as its operation.

5.1. Database

The CANDR database is the component designed to efficiently capture, store, manage, and retrieve data within the system. This database is the central hub and source for the CANDR dataset: it acts as custodian of the dataset's integrity, providing a robust mechanism for data consistency, accuracy, and durability. The database also plays a pivotal role in streamlining the interactions between various other components of CANDR. In a complex system where multiple modules, services, and interfaces coexist, the database serves as the universal intermediary by facilitating data exchanges and maintaining a coherent data model that all other components can rely upon. Finally, through the interface of the website server, the database performs essential CRUD operations, allowing the dataset to be mutated under strict access controls.

5.1.1. Database technologies

I identified three main database technologies suitable for this role: flat file, RDBMS, and NoSQL databases. A "flat file" architecture is where every record or set of records is stored as an independent file in the filesystem. This is commonly used in sites such as blogs where each post or page is a separate file, and a web server converts each file into a page either before publishing (a Static Site Generator) or at every page load (a flat file Content Management System) (Diaz 2018). For example, the *xy*-coordinates of the elements on a stave could be saved as a CSV file in a directory structure with the ID number of the stave as its filename. Erickson's (1970) study of the repertory could arguably be seen as a flat file database on IBM punch cards. This has advantages in the sense that the extra administration of a database is offloaded to the filesystem, and any datatype can be easily stored and indexed, providing it has the correct file extension. This works especially well for images which are already large files and would not require further processing to place within the database. The system could also rely on already existing filesystem backup, versioning, and data integrity tools: for example, a backup could be as simple as creating a compressed archive copy of the directory structure and storing it elsewhere.

The greatest disadvantage to this technology is the requirement for creating and maintaining the changing relationships between elements. Filesystems excel at efficiently storing large quantities of static data; however, the task of consistently linking disparate pieces of data together, especially when dealing with sizeable image and video files, often becomes challenging and necessitates the prior steps of opening, validating, and parsing both files. Moreover, there is no technology ecosystem surrounding this method: a flat file architecture requires the developer to manually ensure consistency, creating their own schema and constraints, or to develop bespoke tools and workflows to ensure consistency.

Moving on from flat file databases, the most conventional database technology in use today is an RDBMS. This is the most well-supported and consistent technology, having first been described over half a century ago (Codd 1970), and as of writing is the most popular database model in use (DB-Engines 2025). As such, there are numerous popular, open source, and well supported RDBMS (e.g. *MariaDB*, *PostgreSQL*, *SQLite*, etc.). In an RDBMS the database is arranged into several inter-related tables. The relational model is a good fit for datasets with a fixed set of columns and where the attributes for each item are known ahead of time (Schreiner et al 2019). The simple data model of an RDBMS allows the database to be fast and efficient for well-indexed data. The ecosystem surrounding mature relational systems are well-suited for the simple data integrity and backup requirements of the CANDR system, and so off-the-shelf tools could be used to implement this. Most RDBMS are ACID (Atomicity, Consistency, Isolation, Durability) compliant, and therefore have certain integrity features bult in (Haerder and Reuter 1983).

A drawback to using an RDBMS is the way in which it interfaces with other components. The most popular and well-supported systems use SQL (Structured Query Language), a query language specific to relational databases. Each RDBMS uses its own dialect of SQL, and every query to the database must first be expressed using parameterised SQL expressions. Database queries must therefore be carefully crafted, as using the wrong query method can be disastrous for performance, data integrity, or even security. The integrity and security issue is well known and comes from failing to properly sanitise data input, allowing an attacker to alter database contents by inserting queries into data fields (Edmunds 2016, pp.1–4).[1] For performance, retrieving multiple items from the database can lead to the "N+1 problem" where retrieving n items results in $n + 1$ queries (Yang et al 2018). This can be glossed over by using an ORM, but this too can often simply hide or even exacerbate the inefficiency through obscurity. There is an impedance mismatch also, where often object-oriented concepts such as models must be expressed in a query language and constrained within a tabular representation. Additionally, relational databases often rely on a fixed record width, and so it is sometimes difficult to store large variably sized items such as long strings of text, or images.

There are more modern alternatives to relational databases collected under the term "NoSQL", and developed in part as a response to the needs of Big Data applications. Two common forms of NoSQL database considered as potentially suitable for this project were document stores and graph databases. Such systems have many of the same advantages as an RDBMS but can store more unstructured data (Leavitt 2010). Document stores have a flexible schema where each document stored within the database can have its own structure and fields, and typically operate through an extensible and human-readable datatype, such as JSON (JavaScript Object Notation) or XML. These databases also often discard SQL as an intermediary entirely, providing complete native APIs for database querying or using a preferred alternative. For example, *CouchDB* uses JSON to store and retrieve data and provides a complete API over HTTP.[2] Alternatively, *eXist* uses XML tools such as *XPath* and *XQuery*

[1] This common error is so well known, in fact, that it is often informally referred to by the name "Bobby Tables", after a hypothetical joke where a child has been named simply to take advantage of this common issue (<https://xkcd.com/327/>). The joke itself even was part of *Python's* documentation of *SQLite* until it was edited out in 2021 (<https://docs.python.org/3.9/library/sqlite3.html>).

throughout.[3] Document stores are therefore particularly useful for processing unstructured data, such as in cases where the exact fields to be stored are as yet unknown or are likely to change in the near future. However, this means that structuring relationships between entities in the database is more difficult as the schema-less makeup of the data does not ensure that the data is consistent.

Graph databases focus instead on the opposite problem, that of the relationships between entities being free-form rather than attributes or structure. The data model is a collection of nodes with links made between any two nodes. SQL is once again not suitable, so specialised query languages for graph data must be used. For example, graph databases *Neo4j* and *Memgraph* both use a specialised query language called *Cypher* to interact with their data models.[4] Once again however, there is a lack of consistency: whereas document stores do not ensure a consistent structure within documents, graph databases do not ensure consistent relationships between nodes. In this case, we cannot say for certain that all nodes of a particular type will be linked to specific other types of nodes. Nor is it simple to find all nodes that exist within a collection (such as in a document model) as these nodes may be dispersed within the graph itself.

Fundamentally, the key issues that NoSQL databases typically attempt to solve are the problems of unstructured data as well as the demands of horizontal scalability where, rather than having a single database instance as a bottleneck or single point of failure, multiple instances of NoSQL databases can collaborate to eventually agree on the data (Schram and Anderson 2012). NoSQL databases are therefore the ideal solution to specific Big Data problems where globally distributed systems with high uptime requirements must deal with at times millions of concurrent connections and partially unstructured, fuzzy data (Gaspar and Coric 2018, p.41).

The disadvantages to these Big Data databases are apparent in the complexity of the technology. The cost surrounding the administration and maintenance of databases designed with the needs of tech giants in mind is significant: these are not databases that can be placed on a server and expected to run unmaintained for years, instead the sophisticated machinery needed is designed to be run at scale with robust infrastructure backing it up. This immediately highlights the importance of the "You are not Google" mantra, and the fact that even large and popular web applications manage to do well without such technology. The question therefore lies not in the feasibility of implementing the requirements using a NoSQL database, but rather in assessing its justification in light of the project's small scale.

---

[2] <https://docs.couchdb.org/en/stable/intro/tour.html>.

[3] <https://exist-db.org/exist/apps/doc/xquery>.

[4] <https://neo4j.com/docs/getting-started/cypher-intro/>, <https://memgraph.com/docs/querying/differences-in-cypher-implementations>.

It is for the above reasons that I chose to use a conventional RDBMS in the design of this project, as it is a paradigm which has numerous well-supported technology options and each with broad support for interacting with other components of the system. The flat file paradigm, although compelling due to its simplicity and perhaps even performant enough for the needs of this project, would require the creation of many bespoke tools to interact and interface with its data. At the same time, there are only a handful of entities that require modelling within the database system (e.g. staves and notational elements) and the requirements of the project do not currently require scaling the database to support the activity of more than a few users. Therefore, the horizontal scalability and schema-less advantages of NoSQL databases is not warranted here, as the complexity trade-off for setting up and administering such a database would detract from the development of the small and isolated requirements of the CANDR database.

5.1.2. Database design

At the logical level, the operation of the components of the database can be viewed as the interaction between abstract entities: which entities are linked together, how they are linked, and how many of each type form each link. This is separate from the physical level of the database implementation — in terms of number of tables, types of columns and indices — which will be described in the following section. Fig. 5.1 shows the interactions between entities in the system as an Entity Relationship Diagram (ERD). In the database implementation, each relationship is modelled using the conventional design pattern in RDBMS of foreign keys (Owens 2006, p.61). Many-to-many relationships are modelled by using associative tables, such as in the "System–Facsimile Relation" table that allows facsimiles to contain multiple systems, and systems to appear on multiple facsimiles. Polymorphic relationships are modelled by using a separate column to indicate which table the foreign key belongs to. The description of the database entities and attributes in this dissertation will maintain the logical level, as the physical implementation of this system is documented in the codebase itself.[5]

The layout of the database follows and extends the schemata outlined in Figs. 4.2 and 4.3 in greater technical depth. There is a broad hierarchy in the first-level schema, of each schema item containing multiple instances of items lower in the hierarchy. This was described in detail in "4.5. Ontology", and is represented in the ERD by a set of one-to-many relationships between the entities of "source", "folio", "facsimile", "system", and "stave". The second-level schema is composed of the entities that represent notational elements on a stave, generalised by a polymorphic relationship into

[5] <https://gitlab.com/candr1/candr/-/tree/master/resources/db/migrations>.

ERD graph of CANDR

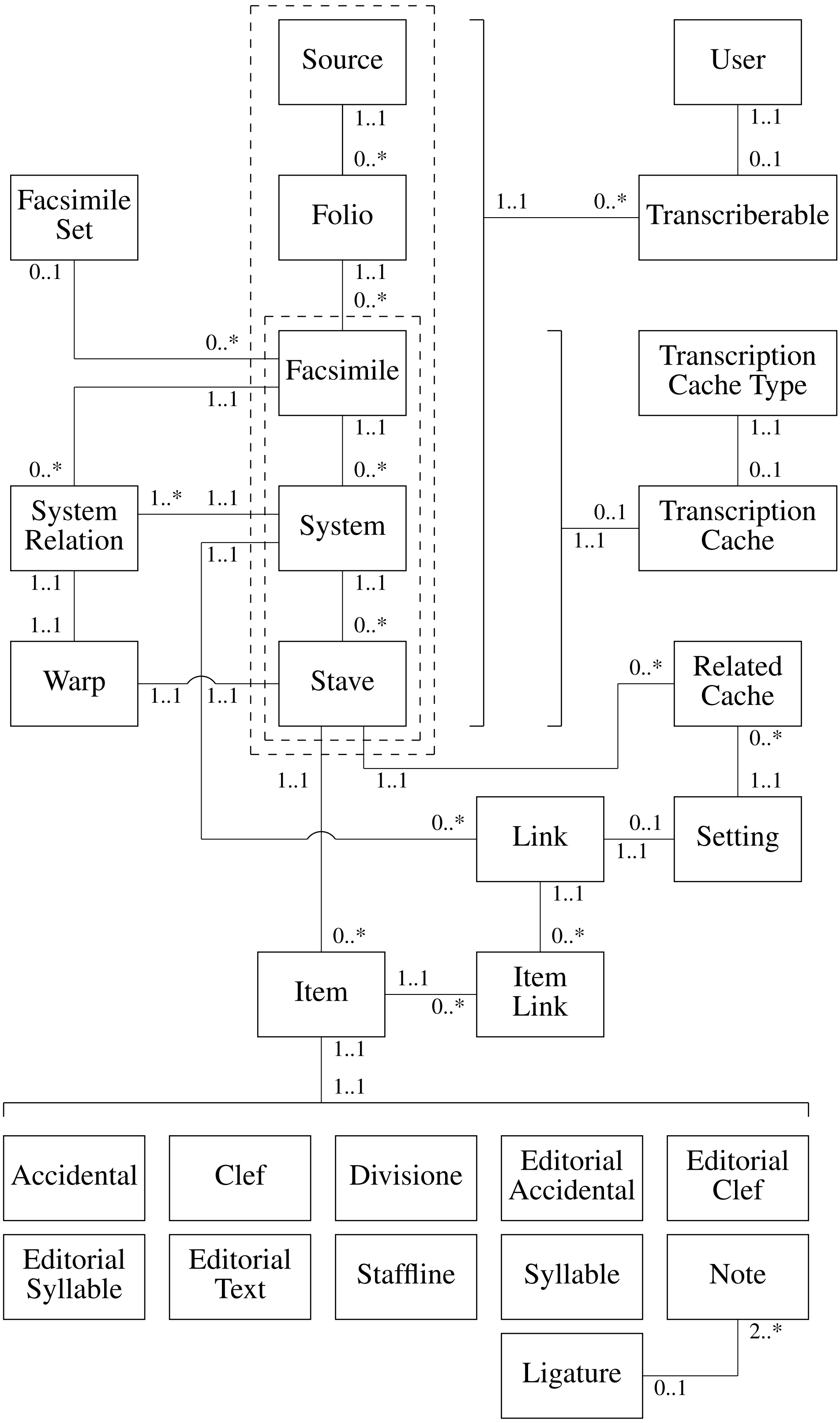

Fig. 5.1

an “item” entity. This is necessary because each notational element must be able to be addressed by a unique ID so that they can be gathered together onto staves.

Each entity in the first-level schema contains two attributes in common. The first is a “transcribed” attribute that is by default set to false, indicating that the data contained therein is either missing or incomplete, and cannot be used reliably either for transcriptions or for ground truth data for OMR.[6] Once a first-level schema entity has been transcribed, the attribute can be set to true and safely used. Secondly, there is also a “history” attribute that extends the data integrity of the CANDR system by encoding an archive of the instance’s previous states, versioned by time.

At the top of the first-level schema is the “source” entity. A “source” can have a “name” attribute that identifies it, which can be any string of characters, and does not have to be unique. No other known and relevant external system catalogues the same granularity of data as CANDR, and so there are few opportunities to create linked data. One possible place, however, is in the “source” entity: the archives that sources are physically located in are all described in linked data resources such as *DBPedia* and *WikiData*.[7] A source therefore does not encode any attributes about its archive location other than an RDF link that provides rich data about the archive. For example, $\mathbf{W_1}$ is located in the Herzog August Bibliothek, Wolfenbüttel. Rather than encoding redundant information about the library in CANDR, the “source” entity links directly to its entry in *DBPedia*,[8] and CANDR can then extract information such as the name and URI of the library for display to the user.

Using linked data in this context offers several advantages. It efficiently integrates data from other sources into the ontology of CANDR, reducing the redundancy and overhead of adding these details manually. This enriches the data that CANDR can draw upon by tapping into external linked data resources and enhancing the contextual information that annotates the data (Antoniou and Harmelen 2004). Adding CANDR into the linked data ecosystem also allows other resources to link back into CANDR and, in turn, enrich their own ontologies through the further granularity of data that the CANDR database provides. CANDR provides unique IDs for every item it catalogues, from sources down to individual notes, and it therefore further provides disambiguation through REST (REpresentational State Transfer) principles (Fielding 2000).

The “folio” entity, linked to the “source” entity, also has a “name” attribute. However, the name of the folio cannot be reliably used to describe the order of folios in a source. Folios may be unnumbered or even mis-numbered, or appear out of order in the source. The CANDR system preserves the original ordering of sources by maintaining an “order” attribute on the “folio” entity,[9] so that folios can be sorted at

---

[6] In the context of machine learning OMR, “ground truth” refers to labelled data that can be reliably used for training, validation, or testing a machine learning model. This is either data that has been input by hand or manually reviewed. See, for example, Fornés et al (2012).
[7] <https://www.dbpedia.org/>, <https://www.wikidata.org/>.
[8] <https://dbpedia.org/data/Herzog_August_Library.rdf>.
[9] Using *Rutorika Sortable* <https://github.com/boxfrommars/rutorika-sortable>.

will.

Linked to each folio there may be multiple facsimiles. The difference for a facsimile over a folio is that it has a related image in the image store. Any time a physical item is to be displayed, it must first find its related facsimile. For example, a folio can be displayed by choosing a related facsimile, and a system instance by correctly transforming its linked facsimile. Simultaneously, groups of facsimiles are gathered together under the "facsimile set" entity. As described in "4.5. Ontology", whole sets of images have been retrieved from various sources. Each have been taken in vastly different ways: camera focal length, colour grading, resolution, to name a few. The "facsimile set" entity contains the common attributes between these images. The most important are those relating to copyright. Many image sources licence their facsimile images under "ShareAlike" licences (Creative Commons 2013). The "facsimile set" entity therefore contains attributes for recording the licencing text to be displayed alongside the image, and a URI to link to the original source. Some licences forbid further sharing of the images — such as "All rights reserved" — and so in this case or in ambiguous cases, an attribute can be set to hide the images to all but authenticated users.

The "facsimile" entity is related to the "system" entity twice. The first way is indirect, through a system–relation association to enable a many-to-many relationship between facsimiles and systems. This is done because not only can many systems exist on a single facsimile, but also a system can exist on multiple facsimiles (consider multiple facsimile sets of the same manuscript). A second way is direct: to set a single facsimile as a system's "transcribing" facsimile so that any time that "system" entity is displayed, it can be easily reasoned which facsimile image should be used.

These relationships, including the connection between the "system" entity and the "stave" entity, are facilitated through the use of the "warp" entity. The "warp" entity plays a pivotal role in defining perspective transforms necessary for cropping and warping sections of the facsimile into systems, and sections of systems into staves. To achieve this, it records the coordinates of the four vertices of the bounding quadrilateral. Whenever a system instance is retrieved, for example, it can seamlessly query the associated "warp" instance in conjunction with the "facsimile" entity to determine the precise transformation required for displaying that specific system. This dynamic transformation approach eliminates the need to store multiple copies of images at full resolution, as images can be cropped and warped in real time by the system as needed.

The "system" and "stave" entities also contain "order" attributes, just like the "folio" entity, to allow their reading order to be controlled by the user. By default this is top-to-bottom but, as described in "1.4.1. Layout and notation", this order is not guaranteed. The "stave" entity also contains an unstructured "predictions" attribute, which contains all the current predictions made by the OMR system (see “5.4. OMR”). When the website client requests predictions for the stave, the server populates the “stave” entity with notational elements from the “predictions” attribute (see “5.3.2. Stave editor”).

The first-level schema entities (“source”, “folio”, “facsimile”, “system”, and “stave”) can be described as “transcriberable”, that is, a neologism indicating that it is possible for them to be edited (transcribed) by users and therefore linked to transcribers (i.e. users).[10] The “user” entity is linked through a many-to-many polymorphic relationship with so-called transcriberables, in that users can transcribe many transcriberables, and transcriberables can be transcribed by multiple users. It is therefore possible for the CANDR system to trace who is responsible for transcribing each entity in the database. User entities themselves represent accounts on the CANDR website. Each is given a user name and password, which is used to authenticate with the CANDR website.[11] Similarly, the entities facsimile, system, and stave can be described as “cacheable”: it is possible to automatically create logical domain transcriptions of these entities from their graphical data, and to save the results in the database for later retrieval (see “5.2.2. Server implementation”). These are related to the transcription cache entity through a polymorphic relationship, so that when a transcription is created, it is saved in the transcription cache. The database also has the functionality for marking transcriptions as stale when a cacheable entity or any one of its sub-entities is altered in the database. Each transcription in the transcription cache is typed by relation to a transcription cache type entity. For example, a transcription may be of type “text” or “MEI”.[12]

Each stave entity may contain many item entities. An item entity is a general entity for any notational element on a stave, such as notes and clefs. Each item entity may have an editorial comment attached as text, and requires a polymorphic one-to-one relationship to specific second-level schema elements. Concordantly, each notational element has a corresponding item entity. Each second-level schema element is encoded exactly as is described in “3.4.1. Graphical domain”, where some elements are encoded as points, some as lines, and some as boxes.

[10] Doubtless, a better terminology than “transcriberable” exists (although “userable” is no better). It is, however, different from “transcribable” which would indicate that it is possible to transcribe these items, not that they can be linked to transcribers. This clunky terminology unfortunately stems from the preferred terminology of the ORM used, *Eloquent*, which prefers the -able suffix to be blithely appended to anything that can be morphed into that thing through a relationship: the documentation includes examples featuring the cumbersome terms “imageable” and “commentable” <https://laravel.com/docs/10.x/eloquent-relationships#polymorphic-relationships>.

[11] In the physical implementation, the password is salted and hashed using *BCrypt*, so that the passwords are not saved by the system, only a one-way hash used to verify logins (Edmunds 2016, p.21).

[12] In practice, the plain “text” type is used for little more than simply demonstrating and debugging the operation of the transcription caches. Regardless, the information within the “text” type is encoded much more richly within the “MEI” type as lyrics aligned to tenor syllables.

One extension to this is in the boxed types accidental, clef, and syllable, as well as their editorial counterparts. Quite often it was found that the box outline of an element on the stave did not map to a pitch, for example the size and shape of a flat is tall and thin, but the logical centre of the flat is at the bottom of the box, not at its geometric centroid. Similarly, the centre of a syllable may be closer to one note, but better aligned with another. The box model in these cases is extended with a freely-moving "centrepoint" that indicates the logical centre of that instance. At the same time, notes can be close to one stave line but should be transcribed at the pitch of another. To cater for this, a "shift" attribute is introduced for the "note" entity that allows an editor to manually correct the automatically-transcribed pitch of a note up or down. The direction of any *plicae* attached to notes are also catered for with another attribute introduced for the "note" entity as an integer indicating the direction and step of the element.

Finally, instances of the "item" entity can be linked together as synchronous using the "link" entity. The "link" entity has a many-to-many relationship to the "item" entity (as items can form part of many links, and links can have many items) using the "item–link" entity. Certain links are tagged as the beginning of musical settings, using the "setting" entity. The "setting" entity has a "name" attribute which is typically given the text incipit of that setting. For purposes of transcription, a setting needs to know which systems it links to, as the setting is only a tag to where it begins and does not know where it ends. This is achieved through another many-to-many relationship using the "related cache" entity.

5.1.3. Database implementation

Turning now to the physical implementation of this design, the CANDR database is implemented as a conventional RDBMS, using *PostgreSQL*.[13] The initial intention was to use the more familiar *MariaDB*,[14] but as many of the data types CANDR requires are long text fields such as JSON encodings of model attributes and large MEI documents, *PostgreSQL* was chosen for its better support for long strings of text. The database itself is complemented by three other sub-components, used for storing images, data integrity, and versioning. These are, in order, the image store, backup, and migrations.

Storing images directly in a database is commonly understood to be bad practice, as they are large files not directly indexable by the database — they must be stored in opaque "blob"-type columns — and requesting images through the database is an unnecessary workload especially when they rarely change. However, images must be tied to an ID in the database for retrieval and linking to images to facsimile records. The common solution to this used in the CANDR system is to maintain a separate image

[13] <https://www.postgresql.org/>.
[14] <https://mariadb.org/>.

store in a directory on disk. Requesting images from the system does not therefore fetch images from the database directly, but the image filename which is fetched and then sent to the user.

The database is backed up by a nightly job into a version control system. This is done by dumping the database contents to disk and committing the changes. The changes are pushed to the *GitLab* service where CANDR's data is hosted.[15] From *GitLab*, the database contents can be reverted to any previous commit, and therefore any past daily state. Having copies in the database, on disk, and off-site therefore fulfils the commonly-used 3–2–1 backup rule.

The database schema, too, is managed with version control. Along with the website server code are packaged database migrations that document using code how the database was created along with the server code.[16] This means that anyone wishing to reproduce CANDR's database can do so by running a single migration command.

5.2. CANDR Website server

The primary function of the CANDR website server (as currently implemented) is to provide an intuitive and user-friendly interface for common CRUD operations on the database. It does this by serving webpages over HTTP to users through a web browser. For more complex operations, such as editing staves, it provides the code for the website client which it then interacts with through an API. It has a separate API for interacting with the programming and analysis toolkit through RDF and MEI.

5.2.1. Server technologies

There are three technologies that were determined to be suitable for this task given the requirements: as an API to be queried directly by the client, as a set of plugins written on top of a general database exploration tool, or as a custom server-side application.

For the first possibility, implementing the server as a thin API over the database has several advantages. Firstly, it allows for a clear separation of concerns between which parts of the application occur on the server and which on the client. The server component can therefore be quite simple as it only needs to wrap the database in a limited set of functionalities.[17] Creating an API also encourages reusability as numerous applications could interface with the same API, i.e. not just the web frontend but the possibility of other platforms such as mobile apps. It is also scalable, as making a clear separation between the client and server allows each component to be scaled

[15] <https://gitlab.com/candr1/candr-data>.
[16] Using *Phinx* <https://phinx.org/>.
[17] The term "wrap" is used in this dissertation to mean a process of encapsulating or enclosing an interface within another class for the purposes of making the interface more compatible with other parts of the system.

independently based on the specific demands of the application. This could be seen as the best, most robust option. However, this would require increased complexity and development effort. The client and server must be very clearly designed in tandem to present a unified interface. The API endpoints must also handle authentication and authorisation, deal with serialisation and validation, and manage the API's versioning and documentation. This is an involved and complex process.

The second technology of writing plugins for an already existing exploration tool is one that would require the least amount of initial effort, as much of the basic functionality of the server and client would already have been created. *Datasette* is such a package for *SQLite* that has begun to see use in academia (Roderic Page 2018),[18] but is primarily a presentation tool and does not support much in the way of editing. *Adminer* is another tool that supports *PostgreSQL* as well as editing, and contains a simple plugin API to extend the functionality of the system.[19] The main advantage of this approach is that the basic web interface and functionality is already present: user login, browsing and editing, exploring the database tables is fully supported out of the box. It should then be possible to write a set of plugins for performing more complex editing operations. However, whether this is truly possible and whether the result would be satisfactory is not known as this has not been attempted before. There may be some hidden technical obstruction in using such a constrained framework, as some of the plugins that would be required would push the limits of what is possible within the system.

Finally, the most conventional technology would be a server-side application using an MVC (Model–View–Controller) architecture. In this design, the server is a larger application that controls all of the business logic and outputs HTML webpages directly to the user. The greatest advantage to this approach is simplicity and familiarity to me as a developer. The MVC pattern is a well established and widely adopted architectural pattern (Sunardi and Suharjito 2019), and it is an architecture with which I already had experience. The database is mediated entirely through the web application, and therefore the security of the system is increased as misconfiguration could not trigger any erroneous results. Only correct outcomes, already programmed into the system, are allowed. The main challenges of an MVC architecture are limited interactivity and scalability. Each click on the application triggers a full page reload, requiring the server to re-render the entire page. This means that any client-side interactivity without full page reloads must go down a different route — such as an API using AJAX (Asynchronous JavaScript and XML) as detailed in "5.3. Website client" — and the monolithic application must be vertically scaled for high demands.

[18] <https://datasette.io/>.
[19] <https://www.adminer.org/en/plugins/>.

5.2.2. Server implementation

The server is implemented as an MVC architecture, using the *Slim* micro-framework in the PHP programming language.[20] The *Slim* framework is similar to larger, more fully-featured frameworks such as *Laravel*,[21] but does not include any of the features deemed unnecessary to this project. *Slim* by default only provides a bare-bones MVC framework, which was in this case extended by using *Laravel's* ORM, *Eloquent*.[22]

A model file was written for each of the database tables to define their relationships and attributes in a way that the ORM can understand. For example, the model file for a *Stave* informs the ORM of the format and types of columns in the table (in this case three timestamps for version control, a zoom level for viewing, a list of predictions from the OMR system, and a boolean flag indicating whether it had been successfully input) as well as its relationships to other models (its containing *System*, the related *Warp* for perspective transform, the *Item*s of notational elements, the *User*s who have contributed to its input, and the *Setting*s that are transmitted here). There are also associated helper functions, for example a function for propagating the message to invalidate cached transcriptions if the model has been altered in the database.[23] The stave's handling of the message finds all settings that contain this stave and sends them the same message. Finally, there is also a data consistency function executed on system start that ensures if the model is deleted then all items on that stave are also deleted at the same time so that they are not left orphaned in the database.

The views are created using the *Twig* templating language.[24] A set of reusable components is created within blocks and macros, such that most views are simple structures of reusable components. For example, the view for editing a facsimile's containing folio extends a generic template for editing model containers and need only alter the title of the page and the names of the models. The HTML is styled with the aid of the *Tailwind* set of CSS classes.[25]

Each first-level schema ORM model generally has two controllers: a "browse" controller and an "edit" controller. The browse controller contains a function to list the ORM model (i.e. fetch all the models of this type from the database and render a list view for this model) and a function to view the ORM model (i.e. retrieve a specific record and render a view to display it). The facsimile, system, stave, and setting controllers also have functions to generate transcriptions. The edit controller contains functions to create and delete records, and two functions for every attribute that the ORM model has: an `edit_*` function and a `set_*` function. The `edit_*` function

[20] <https://www.slimframework.com/>, <https://www.php.net/>.
[21] <https://laravel.com/>.
[22] <https://laravel.com/docs/10.x/eloquent>.
[23] The motivation for this is explained in "4.6. An overview of CANDR" and "5.1.2. Database design".
[24] <https://twig.symfony.com/>.
[25] <https://tailwindcss.com/>.

renders an appropriate editing view for that model attribute, and the `set_*` function receives the result and updates the record in the database.

The most complex controller is the second-level schema *Item* controller. This serves as an interface for the stave editing API. It does not render views as the other routes do, but instead uses a JSON API to communicate directly with the client application. The client stave editor interface (described in "5.3.2. Stave editor") sends a command to this server function, and the server responds with the result. For example, the client can send the command `{"data":{"0":"recall"},"stave":25}` meaning to return all the notational elements on stave with ID #25. The server would respond with a JSON payload with a status code and a list detailing the exact elements and attributes on stave #25. The client can then parse this response and display the result on screen. Other valid commands to the API are: `predict` (for adding OMR predictions to the stave), `delete` (for deleting items from the stave), and `edit` (for updating items). The server is expected to respond with a status code — such as `success` or `error` — and a series of updates for the client to make to the user display.

Finally, there are the logical domain transcribing programs. These are called "walkers" in the server because they are not exactly "transcribers" in the traditional sense of converting to CWMN (see the discussion on p.78), but rather "walk" through the database, automatically inferring a logical domain from the graphical domain plus annotated logical domain paramters (see Fig. 5.2). However, the term "transcriber" is used at times to distinguish their function from other components. An `AbstractWalker` class provides the basic functionality for moving through the database, and concrete classes implement the exact functionality. For example, for any transcriberable item, it is possible to transcribe the syllables using a `TextWalker`. The `AbstractWalker` calls the function `walk()` on every first-level entity in the database, in order, and the `TextWalker` ignores all elements passed to it except *Systems*. The `TextWalker` contains a `TextTranscriber` class, and calls its function `transcribe()` on the lowest stave in the system. The transcriber fetches all the notational elements of the stave, filters by *Syllable*, orders the elements left to right by *x*-coordinate and extracts the text from each, adding a space where the attribute `wordstart` is true.

More complex is an `MEIWalker` / `MEITranscriber`. More than simply syllables, the `MEITranscriber` collects all notational elements from each stave in a system as well as the links between them. For each element on a stave, it draws the stave lines in virtual space and calculates the pitch for each pitched element left to right using the most recent clef and accidental information. It then collects each stave transcription and collates the links between the staves, using PHP's `DOMDocument` class to construct valid XML.[26] Once it is finished, the XML can then be outputted at once. This logical

The hypothetical path of a walker through the database

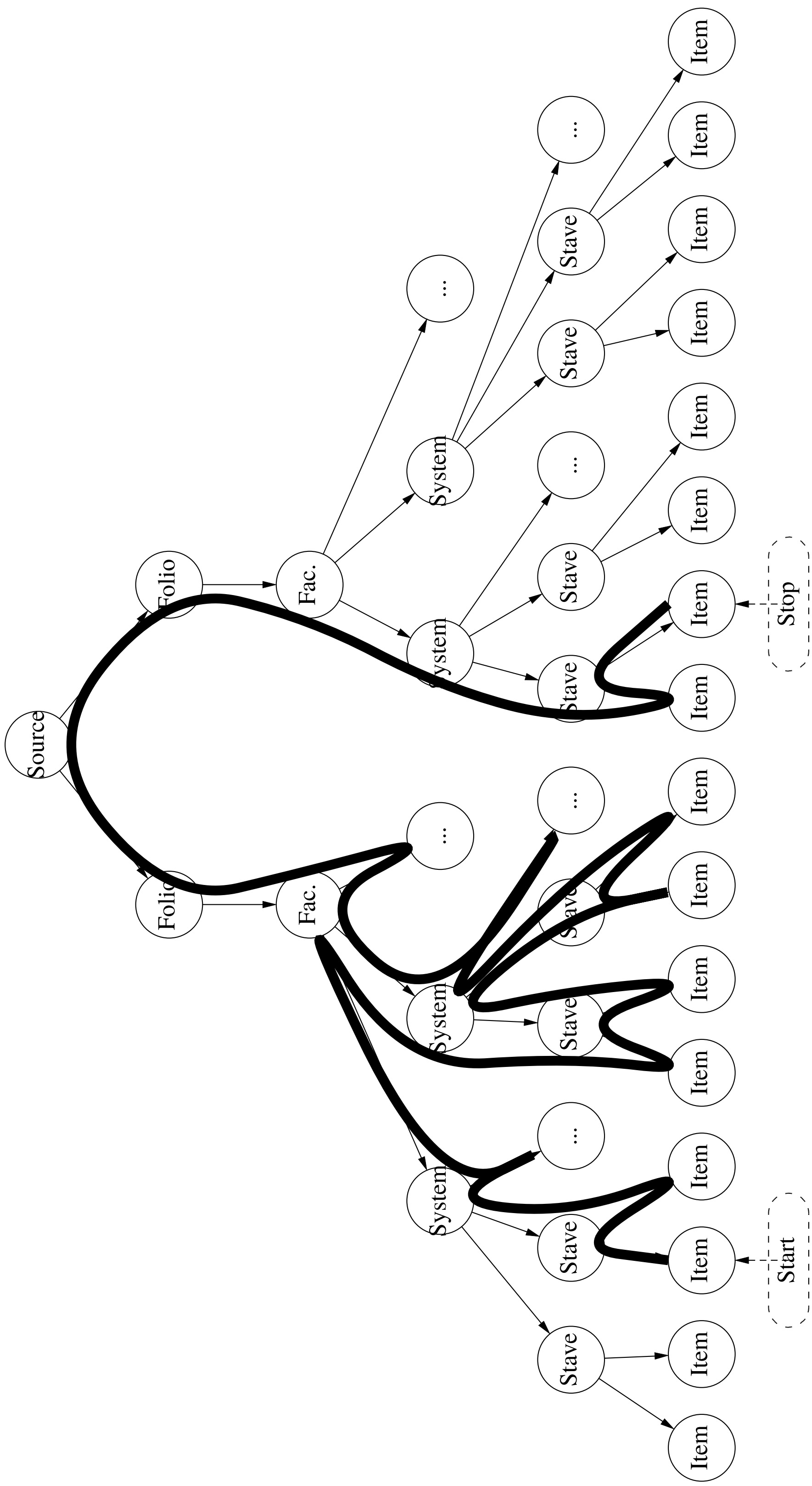

Fig. 5.2

domain music transcription is aided by the helper library *PHPMusicTools*,[27] although

[26] <https://www.php.net/manual/en/class.domdocument.php>.
[27] <https://github.com/benbankes/PHPMusicTools>.

only in terms of pitches and scales, as larger constructions reflect the predominant hierarchical structure (in *PHPMusicTools'* case, inherited from *MusicXML*).[28] This works for physical entities such as systems and facsimiles, but not for virtual entities such as the setting, as a setting is simply a tagged link indicating where the setting begins. Walking a setting is therefore a two-stage process. First a `LinkWalker` walks through the database starting from that setting's link until it finds a link with a tagged setting, or the end of the source. Then, with the start and end position known, these links are passed into the `MEIWalker` / `TextWalker` which returns early as soon as the end link is found.

5.3. Website client

The technology choices for the client are limited by two factors: a) that the only language natively available in web browsers is *JavaScript*,[29] and b) that the client–server interaction has already been set by the boundaries of the server's functionality. It is already known, then, that client functionality is only required for interactions that would not require a full-page refresh, and must be written either in *JavaScript* or a language targeting *JavaScript*. I chose to use *CoffeeScript* as I am already proficient in the language and, at the time of the project's inception, I preferred its Ruby-like syntax over the verbosity of *JavaScript*.[30] Parts of the user interface design were inspired somewhat by the interface layout and workflow of related but non-overlapping projects, particularly the keyboard-first transcription method of *Measuring Polyphony* (Desmond et al 2019–) and the graphical tracing process of *OMMR4All* (Wick and Puppe 2019). There are four pages in the web application that require client-side interactivity, and they each have two modes: browse and edit (depending on whether the user is logged in). These pages are:

1. Facsimile viewer: browsing and editing systems;
2. System viewer: browsing and editing staves;
3. Stave editor: viewing and editing stave elements;
4. System synchronisation editor: viewing and editing links between staves.

One of the key concepts within the client — which is also necessarily replicated in the server transcribers and OMR system — is that of warping the facsimile images to present a straightened, rectangular image. Previous work typically calls this "deskewing", an affine transformation whereby either vertical strips are re-aligned using

[28] The documentation of *PHPMusicTools* states: 'the class structure in PHPMusicTools (PMT) is modelled closely after the markup of MusicXML, and was derived from an older project named PHPMusicXML' (<https://github.com/benbankes/PHPMusicTools/blob/3d31c76368fb9445b1b792af1c8a90e3d2816d40/README.md>).

[29] There is of course *WebAssembly* (<https://webassembly.org/>), but this technology would need to interface with the DOM (Document Object Model) of the browser through *JavaScript* regardless.

[30] <https://coffeescript.org/>.

Editing quadrilaterals, **W**$_2$ f.2v

Fig. 5.3

a heuristic function (Fujinaga 2005), or the straightening of horizontal lines used in text document recognition.[31] I found that affine transformations were unsatisfactory for my purpose as the most frequent imaging issues in the manuscript images were not shears or rotations but rather perspective effects caused by the parchment being bent or the camera orientation not perfectly adjacent to the manuscript surface. An affine transformation cannot help with this as it is only capable of correcting for a 2D transformation. I opted instead for a 3D perspective transform: this was done by re-projecting the bounding boxes (modelled as free quadrilaterals) into the third dimension so that they appear rectangular (Stutter 2021).

5.3.1. Facsimile and system viewer

In browse mode, the facsimile viewer loads the full size facsimile image and displays it on a Web canvas (using *Fabric.js*).[32] It then draws the bounding quadrilaterals for each system defined on the facsimile image, annotates it with its order number, and hyperlinks the quadrilateral to the system's "browse" page. The correct order of the drawing path for the quadrilateral is calculated by sorting the vertices by polar co-ordinate so that they are always drawn anti-clockwise around the first vertex. In edit mode, small handles appear at each of the quadrilaterals' vertices. Clicking and

[31] Many authors use the word "deskew" without writing in detail exactly what the transformation was. The word "skew" indicates an affine transformation. Sometimes the word "dewarp" is used: Wick et al (2019) write that they 'dewarp the line by transforming the image so that the stave lines are straight', but do not write how they accomplished this. The exact methodology behind the image processing is often hidden behind an opaque "pre-processing" step. Conversely, positive examples can be seen in Hankinson et al (2012, p.906) and Thomae (2022, p.104) who both explicitly cite the de-skewing method taken from *OCRopus*, which is affine <https://github.com/ocropus/ocropy/blob/fe78a044691d06f769dafa5876ad552caba36c95/ocropus-nlbin#L71>. Calvo-Zaragoza et al (2020, p.22) use the term "skew-correction" which in my opinion is more amenable to the actual process undergone. Accordingly, here I opt for the active terminology of "warp" rather than the passive "de-warp": this pre-processing step is not returning the image to some original form, but is in fact actively distorting the camera image. To use the prefix "de-" is, in my opinion, to deny the altering impact of the transformation.

[32] <http://fabricjs.com/>.

Stave editor for stave #25, **W**$_2$ f.2r, I

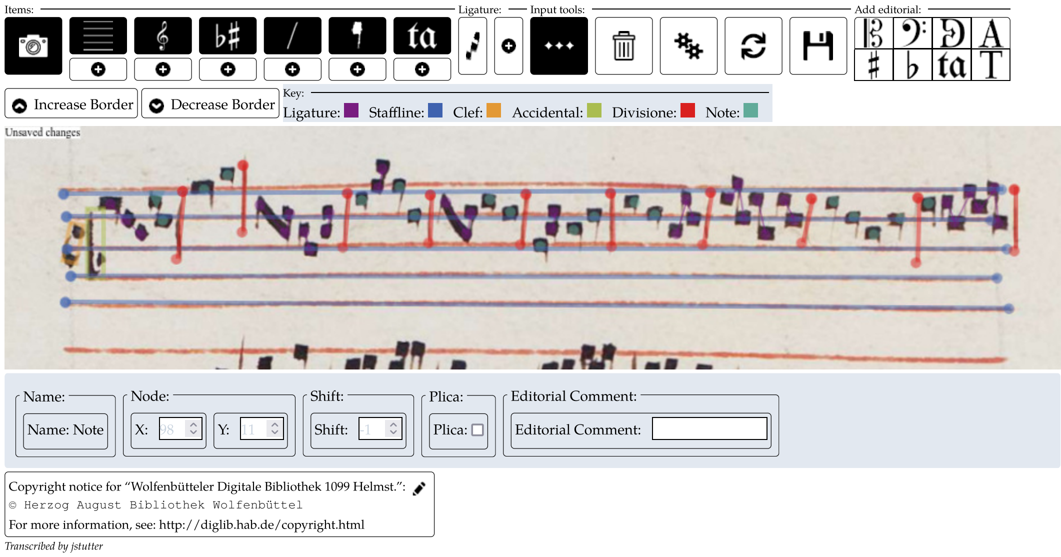

Fig. 5.4

dragging on these handles allows the vertices to be moved, and the quadrilateral is redrawn (see Fig. 5.3). Arrows above and below the order number are buttons that shift the system up and down the transcription order. The facsimile can be zoomed in and out using two buttons in the menu bar, and clicking "Save systems" sends the edited result to be stored in the database. New systems can be added using the "Add new system(s)" button, and the many-to-many relationship between systems and facsimiles can be created by clicking "Add system from other facsimile".

The system viewer is identical, but one layer lower in the first-level schema. However, at this point the system bounding quadrilateral has defined how the system should be displayed to be straightened. The first thing done here, then, is to crop and perspective warp the image using the system's bounding quadrilateral.[33] The system viewer can then present a straightened and warped correction of the facsimile image for that system. These corrections are per-system (and per-stave) so that each system and stave is straightened and corrected to its own parameters.

5.3.2. Stave editor

The stave editor performs the same perspective transform, but twice: once at the system level and then again at the stave level. The canvas is then filled with a straightened stave image. The stave editor is split into two modules: "main" which contains the drawing and editing routines, and "worker" which runs in a background thread and co-ordinates its state with the server. This way, saves do not interrupt the operation of the user

[33] The main logic of this transform was ported into *CoffeeScript* from <https://github.com/jlouthan/perspective-transform>.

interface.

The interface is divided into three sections. From the top down these are: the editing palette, the stave view, and the properties panel. The editing palette itself is split into four panels. From left to right these are: Items, Ligature, Input tools, and Editorial. In browse mode, only the black icon toggles in the Items pane of the editing palette are visible, and the notational elements of the stave are locked in place. However, in edit mode, the view is as in Fig. 5.4.

Beginning from the far left of the editing palette, the camera icon toggles the visibility of the facsimile image itself. The icon of five[34] horizontal lines toggles the visibility of stave lines. The treble clef toggles clefs, the flat and sharp icon toggles accidentals, the slash / toggles *divisiones*, the *virga* toggles notes, and the "ta" syllables. Underneath each of these buttons is a plus icon. This button adds a new element of that class to the stave. On adding a stave line, it appears as a single dot following the cursor which, when clicked by the user into place on the stave, is then locked into place and a line drawn to a second cursor-following dot which is clicked in place to end the line. The same is true for *divisiones*. Notes are placed as single dots. Clefs, accidentals and syllables are modelled as boxes, and these are added at the current cursor position. Clicking on a box brings up multiple handles used to shape the position, size, and angle of the box. A focused box also has a thin line connecting it to a dot indicating its "centrepoint", being the logical centre of the symbol. Clicking and dragging on this centrepoint allows it to be moved around the stave into position.

The small *scandicus* ligature icon creates a new ligature with either the currently-selected note or with the next-selected note if one is not already selected. Clicking the plus icon to the right will add a new note to the ligature.

In the input tools panel, the ellipsis toggle icon is used for bulk insert. When toggled on, the insertion of one item will trigger the creation of a new item of that class without re-clicking on the insert button. For example, the ellipsis toggle allows all the notes to be input at once by clicking the add note button once, and then clicking on all the note locations. The bin icon deletes an element. The button with two gears instructs the server to add the OMR predictions to the stave, if present, and then reload the stave elements. This button disappears once the stave has been marked as transcribed. The reload icon reloads the stave elements from the server, discarding any local changes, and the floppy disk icon forces the worker thread to push an update to the server instantly. The stave autosaves regularly and on exit, so this last button is not required, but is useful for debugging as well as peace of mind. Finally, the

[34] Not to be confused with the so-called "hamburger icon" commonly found on mobile websites to bring up a contextual menu. The "hamburger icon" typically has three lines, this has five and directly represents the usual number of stave lines in a discant stave.

editorial panel adds editorial notational elements. Along the top row are editorial clefs. The D clef has no modern equivalent, and so here I have used the suggestion and design from Hellsten (2016). The bottom row adds editorial accidentals, syllables, and floating editorial text.

A large part of this project is spent "keyboarding", i.e. spending time at the stave editor interface inputting data and OMR corrections. Using the mouse is a slow way of doing this but is the best way of selecting co-ordinates for elements. For everything else, there are keyboard shortcuts to speed up the process: add stave line (L), add clef (C), add accidental (A), add *divisione* (D), add note (N), add syllable (S), new ligature (G), append to ligature (P), bulk insert (M), delete item (Del), increase border (+), decrease border (-), shift note (up arrow / down arrow). I found that input is most efficiently achieved by having the left hand on the keyboard controlling the mode of the editor, and right hand on the mouse inputting the co-ordinates.

At the bottom of the palette are two buttons for increasing and decreasing the border of the stave as, although the image is cropped and transformed by the bounding boxes, sometimes notational elements may continue off into the margins. Next to these buttons are a simple colour-coded key indicating which type of element each colour represents.

Moving onto the stave view, the colour-coded elements are displayed on the stave. The order of ligatures can be understood by the thin lines drawn between notes. It can be at times difficult to ascertain exactly how the positions of notes and clef/accidental centrepoints will be transcribed into pitches. It is calculated in the logical domain transcribing programs by the Euclidean distance from the nearest stave line or space to the point. For ease of use, this part of the transcribing program is replicated in the client. When a note or centrepoint is selected, the stave lines are hidden and replaced with bright green guidelines. Between these guidelines are placed another set of guidelines, equidistant between the stave guidelines, such that every stave line and space has an equivalent guideline. These are all shown as dotted lines except for the closest line to the element, which is shown in solid colour to indicate that the element will be transcribed as that pitch (Fig. 5.5). The centrepoints for syllables operate in a similar way to that of clefs, with the exception that syllables are not attached to pitches by stave lines, but to the notes where that syllable begins. Therefore, when a syllable is selected, it is not the stave lines which are replaced with guidelines in the edit view, but the notes that are replaced by guide points. The closest note to the centrepoint is highlighted in green, indicating that the syllable will be attached to that note (Fig. 5.6). In edit mode, clicking and dragging any one of the handles of an already existing element allows a user to edit the visual parameters of that element.

Just as each entity has a corresponding model in the server ORM, the stave notational elements each have a client model allowing for fine-grained introspection and validation. This is controlled by a `properties()` function which returns a *JavaScript* object of all the properties of the element, each themselves containing an object with `get()` and `set()` functions for that property. The objects also contain type annotations indicating the type of validation to be done on that property. For example, the centrepoint properties object of the clef object is:

```
centrepoint: {
    get: =>
        {
            x: (@centrepoint.getx() * zoom)
            y: (@centrepoint.gety() * zoom)
        }
    set: (dat) =>
        @centrepoint.setpos (dat.x / zoom), (dat.y / zoom)
    type: {
        x: 'int'
        y: 'int'
    }
}
```

When an object is selected in the stave view, the properties panel calls the object's `properties()` function for that item, parses the result, and displays an editable widget for each property. This means that not only can non-visual parameters such as shift or stanza be added this way, but visual parameters such as *x*- and *y*-coordinate can be fine-tuned.

The backend of the stave editor is contained in the "worker" module. When the main thread creates a change to the state of the stave, it pushes a message into a queue for the worker module. Every five seconds, the worker thread wakes up and consumes its queue, compressing the change messages into one large change. If no new messages arrive in the next five seconds, then it packages the messages and makes an AJAX request to the server API (see "5.2.2. Server implementation"). It parses the server response and instructs the "main" module exactly how to update its models to reflect the server data. If there are no new messages in the queue then it then goes back to sleep for the next five seconds. This is done to reduce the number of network calls made. If every change resulted in a network call, then the server could easily be overwhelmed, or race conditions cause messages to arrive out-of-order, compromising the database's integrity. The exceptions to this are explicit save, recall or predict messages which force the queue to be compressed and make a network call instantly.

5.3.3. System synchronisation editor

The system synchronisation editor is fed the links and items of all the staves within the systems as well as the stave bounding quadrilaterals. It would not make sense to

Stave #25, guidelines

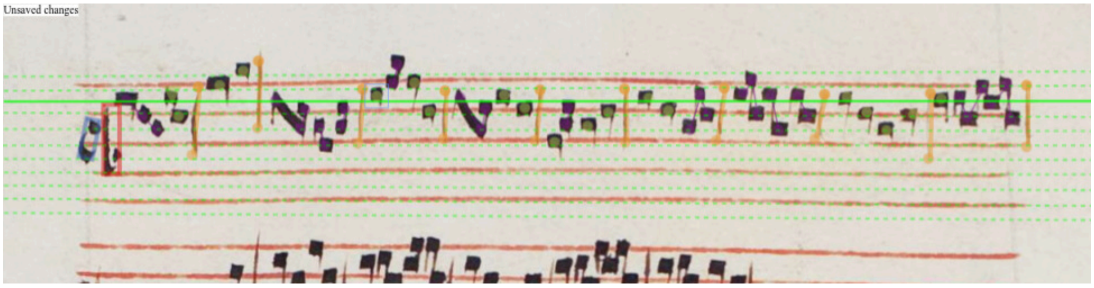

Fig. 5.5

Stave #32, syllable highlighting

Fig. 5.6

synchronise stave lines as they do not represent musical events, and so these elements are filtered out (see Fig. 4.14).

The remaining stave elements are specified in the database by their co-ordinates relative to the top left of the stave and the system synchronisation editor displays all staves for the system, with their elements, in order. However, each stave has its own bounding quadrilateral and therefore its own warp. The system synchronisation editor calculates the inverse of the warp for each stave and transforms the positions of the stave elements to a position relative to the system. They are then added to the canvas as translucent items.

In edit mode, clicking an item highlights it in bright green. Clicking on another item in another stave then draws a line between the two items. In settings with more than two voices, the linked item can then itself be selected to link to another item, and so on. The interface shows an error if a user attempts to link two items on the same stave, either directly or indirectly through an intermediary item, or creates a cycle of links. Clicking on an already-existing link deletes it. The links can then be updated in the database by clicking on “Save links”. Clicking on “Add Setting” prompts the user to select a link on the system. Selecting a link creates a new setting in the database and brings the user to a page for editing that setting. A link cannot be used for more than one setting.

5.4. OMR

Fujinaga (1996) coined the term “adaptive” OMR to apply to an OMR system that ‘will

Workflow of adaptive OMR from Fujinaga et al (1991, p.66)

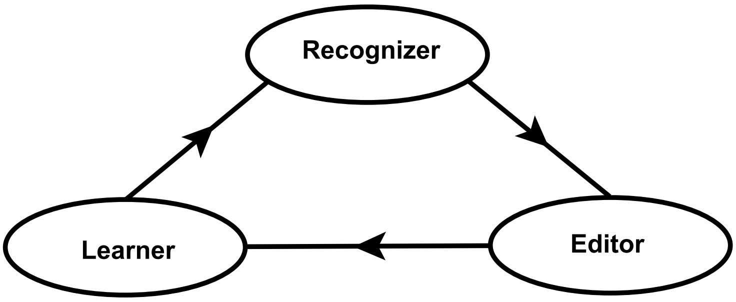

Fig. 5.7

continually improve its accuracy' (Fujinaga 1996, p.2). In such a system, the human is kept in the loop as an editor to continuously correct and guide the OMR training process. This differs from a more typical process where a subset of the data is manually processed and used as ground truth for a one-shot training and prediction process. In adaptive OMR, although there is an initial phase of manual processing, a constant loop of prediction and correction gradually improves the OMR results (see Fig. 5.7).

CANDR's OMR system is added into the loop in this same adaptive way. The OMR is not the keystone of the project, but simply a method of accelerating an otherwise fully manual keyboarding process. The OMR therefore does not need to be state-of-the-art, but simply workable and "good enough" to provide a significant advantage over manual entry (Stutter 2021). There is also an element of restraint that had to be factored in here: OMR is yet another avenue down which the research could easily be distracted from this project's core aims. The role of OMR in CANDR is described below but not considered wholly novel, although it is an area for further investigation and development.

5.4.1. OMR technologies

Fujinaga et al (1991) set out the most common modern workflow for adaptive OMR: pre-processing, segmentation, feature extraction, and classification. The process has been refined in the intervening 30 years and the individual sections have been expanded or altered, but this core methodology has remained the same (Bainbridge and Bell 2001). In this methodology, images of staves or even whole facsimiles are passed into a series of machine learning processes. There is usually some element of initial pre-processing performed on the images to make them suitable for machine learning, such as rescaling, de-noising, and warping. After this pre-processing, one of the first processes in this methodology is to detect stave lines and therefore remove them, so that the intersection between stave lines and musical symbols does not confuse a classifier

(Rebelo, Fujinaga et al 2012). From there, the symbols can be segmented, features extracted and then classified. It is important to note, however, that this workflow was created with CWMN in mind and many of its assumptions — such as thin, straight stave lines and well-spaced notation with few overlaps — do not hold true for historical notations such as ND polyphony. Nevertheless, there has been some success using this method in printed music or even in regular chant notation (Hankinson et al 2012).

For handwritten and early music, a different workflow is often required. A pioneering early example can be found in Pugin (2006), who used a Hidden Markov Model (HMM) on slices of the stave directly. Critically, many of the pre-processing and segmentation steps can be skipped: the staves need not be perfectly level, and 'the features are extracted directly from the image without erasing the staff lines first' (Pugin 2006). On early music prints, this yields better results than segmentation models (Pugin et al 2008).

The current state of the art has now 'caused some steps to become obsolete or collapse into a single (bigger) stage' (Calvo-Zaragoza, Hajič Jr, and Pacha 2020, p.23), and this has been brought about by the advent of Deep learning models now used as end-to-end OMR solutions. There has recently been an explosion of literature based on the training and evaluation of various architectures for Deep learning OMR in the last ten years. Whereas previous machine learning approaches to OMR used techniques such as SVMs and *k*-nearest neighbours (*k*-NN) to classify stave elements — see, for example, Rebelo, Tkaczuk et al (2011) — Calvo-Zaragoza et al (2017a) used a CNN for stave line detection and removal, and a second for stave elements. Calvo-Zaragoza et al (2017b) used a CNN based on the VGG (Visual Geometry Group) network for pixelwise classification without removing stave lines. Hajič Jr. et al (2018) used a U-Net topology trained on the convex hulls of individual symbols. Calvo-Zaragoza and Rizo (2018) used a CNN followed by a Long short-term memory (LSTM) recurrent neural network (RNN) to create output as a sequence of symbols. Similarly, Baró et al (2019) used a CNN in sequence with an RNN for sequence prediction.

Generally, a CNN is more suited for image detection and segmentation: creating heatmaps of feature probabilities for further processing. RNNs are more often used to process sequential data, for example reconstructing the feature outputs of a CNN into a logical series of symbols. More recently, work has been done to improve the accuracy and performance of these models, by applying post-processing or altering the models slightly. Calvo-Zaragoza et al (2019) used an HMM in tandem with a multilayer perceptron (MLP) to improve accuracy over a simple HMM. Paul et al (2022) used an ensemble of different Deep learning methods to improve overall accuracy. Wen and Zhu (2022) applied the transformer architecture to the output of a CNN and LSTM to improve accuracy over simpler models. This only represents a small slice of recent research in this area, and it is a rapidly evolving field.

OMR systems have also been placed as key components in end-to-end music document input systems. The archetype for this is *Rodan* (Hankinson 2014),[35] but there is also *MuRET* (Rizo et al 2018),[36] and *OMMR4All* (Wick and Puppe 2019).[37] Unfortunately, each of these systems is tightly coupled to a particular repertory, process, or ontology of music that is not applicable here. For example, *Rodan* is created for large, distributed systems and for applying a similar set of transformations in multiple different configurations. *MuRET* is, like *Rodan*, tightly coupled to MEI and *Humdrum* and therefore neither system has support for editing ND polyphony in the graphical and graph-based format that is pertinent to this study. *OMMR4All* only supports monophonic music or partbook layout, and does not support the ability to edit or align music in score format. No models for OMR of ND polyphony already exist and so, even if I were to use an already existing system, I would still likely have to train my own model to work with that system as the consistent application of transfer learning to OMR is still in its infancy (Baró, Riba, Calvo-Zaragoza, and Fornés 2019).

There are two alternatives to these systems. The first is to exercise the aforementioned restraint, and decline to explore the possibility of OMR. The CANDR editor described in "5.3.2. Stave editor" is already sufficient for the requirements of the project, as without OMR it is still possible to enter the requisite data. Given the above facts about the state of OMR development to date, it may not be possible within the constraints of the project to create an OMR system accurate enough on this notation type, replete as it is with ligatures and scribal idiosyncrasies, and the time saved may in fact be very small. Worse, the time spent implementing an OMR system for this notation may eat up more time than it is worth, and the time, effort, and resources investing in developing and integrating OMR technology may outweigh any potential time savings, resulting in a net loss of productivity.

The second alternative is to implement an OMR system from scratch for this exact purpose. Consumer-focused machine learning frameworks have made the implementation of Deep learning networks much simpler, even within the last few years. Implementing a custom OMR solution for CANDR would allow me to design the system to my own requirements. Implementing OMR from scratch is still far from trivial however, as it requires an in-depth knowledge of the field of machine learning, and of neural network libraries and topologies, not to mention a high level of programming skill. This difficulty is compounded by the size of the dataset the CANDR system is working with: this is not a toy example of a few pages of manuscript, but the

[35] <https://ddmal.music.mcgill.ca/Rodan/>.
[36] <https://grfia.dlsi.ua.es/hispamus/software.html>.
[37] <https://ommr4all.informatik.uni-wuerzburg.de/en/>.

OMR system would need to train over whole manuscripts at once. This would require a complex pipeline of image retrieval, processing, and streaming as it is unlikely that the entire dataset would fit in memory.

5.4.2. OMR implementation

Notwithstanding the known problems and potential issues, I decided that the implementation of an OMR system was worth the effort required. Given the alternatives, I decided that the best option was to implement an OMR system from scratch, and to do this by training a CNN model to recognise ND notation. Calvo-Zaragoza et al (2017b) introducing their first experiments on using a CNN for layout analysis, carefully remark that:

> We are not trying to find the best network topology — which should involve a comprehensive set of experiments to find the best set of parameters — but to demonstrate that the layout analysis of musical documents based on pixelwise classification with CNN is feasible (Calvo-Zaragoza, Vigliensoni, and Fujinaga 2017b).

Similarly Ríos-Vila et al (2019, p.29) develop an OMR system well 'aware that OMR technologies do not provide perfect results, and they probably never will', and Rizo et al (2018) admit that 'even state-of-the-art OMR systems are not ready yet to operate in an automatic, unsupervised way [...] the intervention of a human expert is a must in the recognition and edition process'. The research currently being undertaken to apply Deep learning to OMR systems is still exploratory and highly experimental, and cannot be "let loose" on a source without constant human guidance and correction. Therefore, even in the most state-of-the-art systems, 'the main goal of OMR is to minimise to [sic] human effort' (Wick and Puppe 2019).

In the same way, this small, ad hoc OMR system as part of CANDR is a first of its kind, and an experiment designed to accelerate and support human efforts rather than to supplant them. As far as I am aware, there have been no attempts to train an OMR model on ND polyphony. State-of-the-art techniques are likely to complicate matters, and so I opted for the simplest technological solution that I believed was most likely to succeed and that best played to my strengths in terms of knowledge of the domain. This study is not interested in a "best" OMR model for ND polyphony — although this is perhaps a direction for future research — but a model that will be useful to the project in that it can be developed quickly and save time in the long run.

The most difficult parts of the OMR system to develop were the pre-processing of the training data, and the architecture for streaming training data from disk. Recall from "5.3. Website client" that the stave items are placed on the stave relative to the co-ordinate system of the stave itself. An element placed at the top-left of a straightened stave will be given the co-ordinate (0, 0), regardless of where the stave is in the system or facsimile. However, the CANDR website server does not provide an API for warping stave images directly: this warping is done in the client. The OMR system therefore needed to first replicate this warping ability in order to generate stave images in the same co-ordinate system as the client would expect to display.

Each source contains different scribal practices, qualities of vellum, image sizes and quality, and subtly different notation. As a result, it was decided to train a separate model for each source. The first step that the OMR system takes is to retrieve the facsimile images and ground truth data (being already transcribed staves) from the database. The warp records are used to warp and crop the images into individual stave images ready for processing in a process identical to that performed by the client. These images are saved into an *SQLite* database on disk.

I wished to minimise any further pre- and post-processing steps, and therefore took a data-oriented approach to design the OMR system. A key part of this was taking a close look at the input data, and exploring ways in which the features could be more easily extracted. A key insight from this approach was that stave lines and stave elements, although both considered equal items in the CANDR system, do not share many features. An obvious observation is that stave lines are red whereas stave elements are black. Beyond that however, in ND notation each stave element is a single connected component: a ligature, note, clef, etc. Stave lines on the other hand are present in nearly every area of the image and connect through the image. They are often interrupted by stave elements overlapping with the line, yet the sense of the line continues. They would therefore require a different set of processes to the smaller-scale stave elements.

I also needed to look at the exact ground truth data that I had — or could make — for training the system. A typical workflow is to use a pixel mask editor such as *Pixel.js* to generate ground truth for several facsimiles, and then use this data to train the OMR model (Saleh et al 2017). This is a laborious process and can take hours for each page, even with the help of heuristic tools such as binarisation and layout models such as those presented in Castellanos et al (2018).[38] The "ground truth" dataset of CANDR is in the graphical domain, just like pixel-based ground truth used in previous work, but here I am not striving for pixel-perfect accuracy. Each of the stave elements are modelled after simple geometric shapes: boxes, lines, points. I therefore experimented with mocking up image masks for each stave automatically using the CANDR dataset. These masks cannot be pixel-accurate as the dataset itself is not pixel-accurate, and the

[38]For example, Thomae (2022, p.90) claimed that her use of the layout model from Castellanos et al (2018) reduced the classification time from 'hours to minutes' — apparently citing the ballpark figures given by Castellanos et al — but it is unclear how this scales to larger sets if each image must still be checked for pixel-perfect accuracy. Missing, too, is the time saved over inputting the data manually.

reality of the notation does not always fit the model (for example, rarely are stave lines perfectly straight). However, the masks could be considered “good enough” for an OMR solution. If these image masks were successful, it would allow the OMR system to bypass the pixel-based ground truth entirely, saving countless hours of work.

This lack of perfect ground truth data also meant that quantitative evaluation of the model was difficult to implement. OMR models for image recognition tasks typically use evaluation metrics such as F-scores to determine the performance of the model: how accurately did the model predict the pixel in test data? However, these metrics assume that the ground truth is perfectly accurate. To ascertain how closely the model matches the masks, the masks themselves would have to be first evaluated against pixel-perfect data, which was exactly the kind of OMR this system was trying to avoid.

It is worth reiterating that the main objective of this tool was not to achieve a pixel-perfect recognition of ND notation but rather to significantly accelerate the input process over manual data entry. In this context, the most important metric for evaluating the system’s performance shifts from precision or recall to the efficiency gained through automation. The primary focus was therefore on quantifying the time saved and the reduction in manual effort achieved by utilising the OMR system. Metrics such as input speed improvement, reduction in data entry errors, or the overall productivity gain became more critical in assessing the success of the tool.

The previous section highlighted recent work that paired a CNN with an LSTM to perform direct symbolic output. Considering once again the musical domains in which this study situates itself, the input for the CANDR system is graphical, not symbolic, and therefore these more modern advances will not be useful. The server, with its logical domain transcription programs, handles the conversion process from the graphical to the logical domain: the OMR system was needed only to create annotated output in the graphical domain. A CNN was all that was necessary, then, to perform what is a fairly typical image recognition task.

Another simplification that was made came from the training of a new model for each facsimile set individually. As every training sample for a model is written in the same manuscript within roughly the same parameters and the same imaging technology, it can therefore be assumed that elements to be recognised will all be of roughly the same size. The use of any image pyramiding techniques commonly used to detect features of different sizes can therefore be discarded as, for example, there are no clefs twice as large as others (Pang and Cao 2019).

I decided to implement the OMR system in *Python* and *TensorFlow*.[39] I selected a small CNN model as a starting point for fast training and as proof that the methodology

[39] <https://www.tensorflow.org/>.

W_2 f.13r III, with CANDR image ground truth image mask for both stave lines and stave elements, from Stutter (2021)

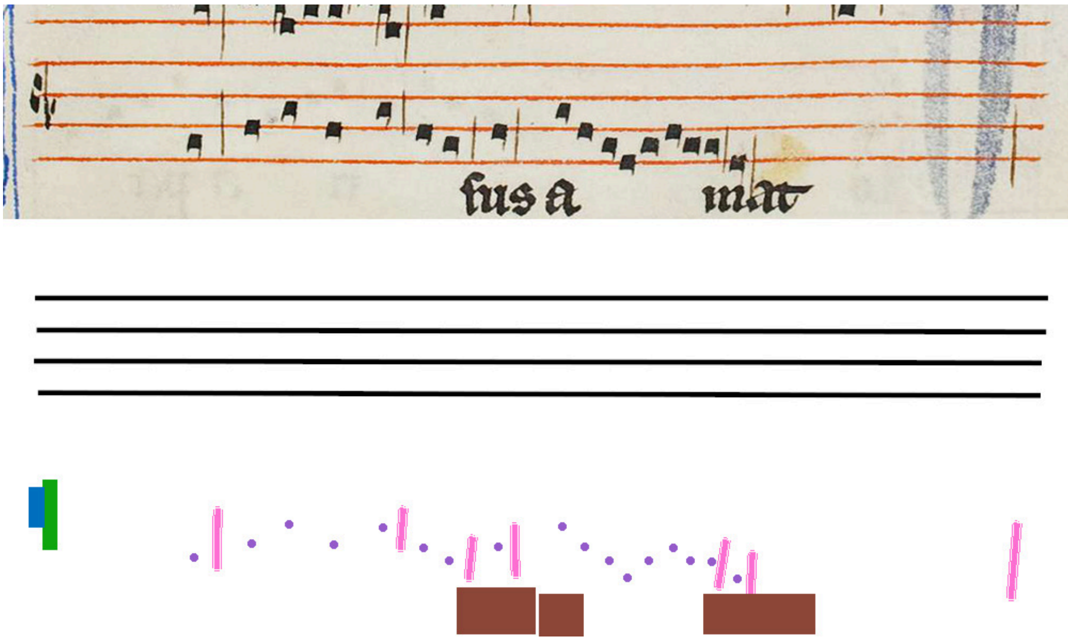

Fig. 5.8

was feasible. This was two sets of 2D convolution layers (3 × 3 with ReLU activation), each followed by 2D maximum pooling (2 × 2). This was then followed by a flattened layer and a 256 dense layer, followed by an output layer. The aim was to train two small models: one for stave line detection, and another for stave element detection.

The training process consisted of a multi-process pipeline in two steps. All the staves in the database were split between several processes — one for each CPU core — and each process generated two masks: a stave line mask and a stave element mask. This was accomplished simply by using the image drawing capabilities of the *Python Imaging Library* (PIL).[40] The mask generation was parameterised such that the attributes of the geometric shapes could be altered to improve the size of their expression in resulting prediction masks. For example, the point-based notes were modelled as circles and the radius of the circle could be altered. The masks were drawn at a larger size and then resized to fit the image (supersampling) as well as to reduce antialiasing artefacts. A small amount of Gaussian blur was also added (see Fig. 5.8). These last two steps were done to allow the network to learn that the masks were not pixel accurate, and that pixels next to an element may be incorrectly classified in the ground truth. The processes were also responsible for warping the stave images themselves. Experiments were done on further image processing, such as converting to greyscale, binarisation, edge detection, but it was found that these in fact hampered the

[40] <https://python-pillow.org/>.

training accuracy.

Then, taking the mask and stave image simultaneously, a sliding square window was manipulated over both the image and mask at a stride length which increased proportionally to the amount of training data available. The most successful window size varied depending on the quality of the facsimile images. To eliminate a stippling effect in resulting predictions created by the uniformity of the stride, the start points of this window were chosen as a random integer between zero and the stride length. For each window, the intensity of the central pixel for each element class was encoded into an output vector in the range 0–1. Therefore, for a set of elements L, the resulting vector was $(L_1, L_2, L_3, \ldots L_n, 1 - \Sigma L)$. For example, in a stave line detector (a one-class detector) where the current position has an intensity of 0.7, the output vector would be (0.7, 0.3). The offsets and intensities were then passed from each process to a queue which stored the results as records in an *SQLite* database. With this pre-processing done, the training process could then be streamlined into a simple process of streaming records from the *SQLite* database into the CNN, which was implemented using *TensorFlow Keras*.[41]

The prediction process for un-transcribed staves was very similar: using multiple processes, every stave was image warped according to its record and split into offsets using a sliding window. Each offset was then passed to a queue to be fed into the CNN, and the result assembled into a heatmap containing layers for each item to be predicted. The stave data was then normalised and binarised using Otsu's method. From here, the post-processing depended on the item to be predicted.

For box-like items such as clefs, connected component analysis was also done on the result and the minimum area rectangle found using *OpenCV*.[42] For lines such as *divisiones* and stave lines, connected component analysis was done and the extremes of *x*- and *y*-coordinates within the component calculated as the extent of the line. Training the stave lines separately allowed the stave line model to effectively "skip over" interruptions from overlaps, and no further post-processing had to be undertaken beyond connected component analysis. Notes were the most difficult to post-process, as ligation often caused multiple notes to become single components. Binary erosion was therefore performed on the heatmap using *SciPy* and the number of connected components recalculated.[43] This process was repeated until there were no more components left. The iteration that yielded the highest number of connected components was then re-selected, and this was used as the basis for constructing notes. Finally, these predictions were packaged up into a single JSON file containing all that

[41] <https://www.tensorflow.org/guide/keras>.
[42] <https://opencv.org/>.
[43] <https://scipy.org/>.

stave's predictions and saved in the database record.

This OMR system was run periodically on each facsimile set as I transcribed it. Due to the large amount of data available from the overlapping windows of each stave, it was sufficient to train the model over the data in a single epoch.[44] To evaluate its performance, I elected to consider how it improved the data input (keyboarding), as more quantitative evaluations were not forthcoming due to the lack of pixel-perfect ground truth data. I found that the most efficient manual workflow was as follows: input stave lines, input clefs and accidentals, input notes, ligate notes, input *divisiones*, input syllables (if present), final corrections. On the other hand, the automated workflow consisted almost entirely of corrections, such as splitting notes and altering bounding boxes. However, the OMR system was only trained to recognise individual notes and so ligation was done by hand.

To measure the time difference per stave, I conducted a small experiment where I input 20 staves manually and 20 using the OMR system. The results of the experiment can be seen in Tbl. 5.1. The mean and standard deviation of the manual sample was 141.08 and 42.66. For the OMR sample, it was 90.46 and 18.40. To evaluate whether the OMR methodology provides a significant reduction in processing times compared to manual transcription, I employed a Bayesian analysis to estimate the difference in mean times and quantify the associated uncertainty. Let y_{man} and y_{omr} be the observed data for the manual and OMR observations, respectively. These are modelled as being drawn from normal distributions:

$$y_{\text{man}} \sim N(\mu_{\text{man}}, \sigma_{\text{man}})$$
$$y_{\text{omr}} \sim N(\mu_{\text{omr}}, \sigma_{\text{omr}})$$

The mean of manual times (μ_{man}) is given by:

$$\mu_{\text{man}} \sim \text{Gamma}(\alpha_{\text{man}}, \beta_{\text{man}})$$

where:

$$\alpha_{\text{man}} = 5,\ \beta_{\text{man}} = \frac{\alpha_{\text{man}}}{\bar{y}_{\text{man}}}$$

and $\bar{y}_{\text{man}}$ is the sample mean of y_{man}. The mean of OMR times (μ_{omr}) is given by:

$$\mu_{\text{omr}} \sim \text{Gamma}(\alpha_{\text{omr}}, \beta_{\text{omr}})$$

where:

[44] "Epoch" is a term used in machine learning to mean one pass over the training data. For example, a model trained in ten epochs sees the training data ten times.
[45] <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.normaltest.html>.

$$\alpha_{\text{omr}} = 5,\ \beta_{\text{omr}} = \frac{\alpha_{\text{omr}}}{\bar{y}_{\text{omr}}}$$

and $\bar{y}_{\text{omr}}$ is the sample mean of y_{omr}. The standard deviations (σ_{man} and σ_{omr}) are given by:

$$\sigma_{\text{man}} \sim \text{HalfNormal}(\sigma = 100)$$
$$\sigma_{\text{omr}} \sim \text{HalfNormal}(\sigma = 50)$$

The difference in means (Δμ) is defined as:

$$\Delta\mu = \mu_{\text{man}} - \mu_{\text{omr}}$$

and the posterior distribution:

$$P(\mu_{\text{man}}, \sigma_{\text{man}}, \mu_{\text{omr}}, \sigma_{\text{omr}}, \Delta\mu \mid y_{\text{man}}, y_{\text{omr}})$$

Samples were drawn from this posterior using Markov Chain Monte Carlo (MCMC) with 2,000 iterations. The posterior mean of the difference in means (Δμ) was 50 seconds, with a 95% HDI (Highest Density Interval) of [28, 72]. This means that, with 95% probability, the true improvement in mean transcription times using OMR lies between 28 and 72 seconds, with manual processing consistently taking longer than OMR.

Development of the OMR system was successful as it increased the speed of data input by over a third. The effect of this was not just physical but also perceptual: the workflow of correcting OMR predictions felt like a faster and cleaner process, as there was less repetitious work being done, and even if some elements had to be corrected, it felt like a smaller effort than beginning each stave from nothing.

It was mentioned at the opening of this section that the OMR system was not designed as an end goal, but rather as a method for improving the speed of an already manual process. This it has done, but it is clear that the system could be greatly improved to reduce the amount of human correction required. This OMR system only considered a small and shallow CNN, and larger and more complex network topologies such as VGG, R-CNN, and YOLO could be harnessed to improve the results (Simonyan and Zisserman 2015). Moreover, given more time, the model could be evaluated more robustly, for example against a random classifier, or a model trained on CWMN or plainchant. Nevertheless, this model can provide a baseline result for future models to evaluate against.

5.5. Conclusion

This chapter has presented the methodologies behind the “online” aspects of the

CANDR system: database, website server, website client, and OMR system. For each aspect, available technologies have been discussed and evaluated, and the exact implementation of the system described. For certain conventional parts of the system such as the database and website server, only the minimum required to explain the exact operation of the system has been included. However, for certain other aspects — such as the transforms of the website client and the creation of the OMR system — some parts of the methodologies are novel or not yet applied to this kind of data. These have therefore been described in greater detail so that the intricacies of their operation can be fully understood. The analysis and selection of suitable alternatives ensured the chosen technologies were appropriate to the system's requirements. There is one final part of the CANDR system to be described — the programming toolkit — and its methodology is completely novel. The next chapter will therefore be devoted to a description of its operation, and the analysis approaches taken in this study.

Results of experiment measuring
manual vs OMR-assisted input

Sample #	Manual (s)	OMR (s)
1	150.15	92.44
2	134.50	85.28
3	206.17	86.83
4	114.50	79.24
5	162.51	75.32
6	164.72	101.74
7	104.23	84.73
8	192.66	91.86
9	84.74	103.70
10	123.51	113.66
11	136.02	69.05
12	90.78	118.60
13	192.13	90.80
14	153.38	78.94
15	125.49	67.51
16	94.27	119.30
17	201.78	105.32
18	83.56	76.12
19	96.83	115.37
20	207.80	53.43
μ =	**141.08**	**90.46**
σ =	**42.66**	**18.40**

Tbl. 5.1

Chapter 6
Offline

The methodologies described in the previous chapter were concerned with the creation of a dataset of ND polyphony suitable for analysis. While this dataset is a vital prerequisite for advancing research into the ND repertory through digital methods and substantiating the research at hand, it does not, on its own, address the central research question regarding the detection and analysis of musical reuse within the broader context of the ND repertory. Patterns of musical reuse may be dispersed across the repertory, and so the analysis required must be capable of considering the corpus as a whole. To bridge this gap, this chapter extends the foundations laid out in "Chapter 3: Data models", and charts a course towards methodologies that can be applied to the CANDR dataset to provide evidence of musical reuse patterns embedded across the ND repertory. In essence, this chapter offers a response to the problems posed in "3.6.4. Encoding conclusion" and "3.8. Methodological utility", namely: how best to encode the logical domain of ND polyphony, and what analytical methodologies can reliably be used. These issues will lead directly to detecting musical reuse in the ND repertory.

"3.8. Methodological utility" posited an ideal methodology that would serve to create evidence for the research questions for an equivalent monophonic corpus: this was 'to encode small sections (such as ever-increasing n-grams or skip-grams) in an embedding, mapping these n-grams into a d-dimensional space, and clustering similar items together' (p.117). This form of methodology has only recently become more common, but has already seen success in monophony: in extracting common patterns from Arab-Andalusian music (Nuttall et al 2019), as well as plainchant (Cornelissen et al 2020). Attempts have been made to map this same word-embedding methodology into CWMN, by "chordifying" polyphonic music into discrete vertical slices (Chuan et al 2020)[1] but, as was seen in "3.7. Polyphonic problems", this technique collapses the important counterpoint between the individual voices into a series of time slices and, as it stands, cannot therefore reasonably be performed for ND notation.

As I see it, there are three methodological obstacles to be surmounted in order to be able to apply embedding methodologies to polyphonic music in a way that captures at once both the sequence and alignment of ND polyphony: a) the method of modelling (how the music is represented in the computer); b) the method of embedding (creating a

[1] This "chordifying" terminology comes from the music analysis framework for *Python*, *music21* <https://web.mit.edu/music21/>. Also called "salami-slicing". See Sears and Widmer (2020) and Cuthbert and Ariza (2010).

vector for each token); and finally c) the method of clustering (pulling similar items together). This chapter will present workable solutions to each of these problems in turn, and then present the programming toolkit created to implement these techniques and synthesise them into a cohesive methodology.

6.1. Modelling

"3.7.1. N-grams" highlighted the usefulness of n-grams and skip-grams for creating collocations in a corpus. This methodological underpinning is so effective as to become impossible to ignore for the purposes here. Text can be tokenised into the smallest unit deemed useful: this could be a letter or word. An earlier example tokenised two Brahms themes into a vocabulary of pitches (Figs. 3.21 and 3.22). However, for polyphonic music what it means to "tokenise" becomes unclear as, at the very least in ND notation, it cannot be said with certainty which notes align with which.

"3.7. Polyphonic problems" found that there are typically two methods for tokenisation and modelling used for analysing polyphonic music: considering each part individually as monophonic and measuring the steps between pitches per part, or considering vertical slices ("chordifying") along a time axis (Cumming et al 2018).[2] This dual axis multidimensional analysis directly reflects the two axes of tabular music representations outlined in "3.6.1. Tables" and cannot be used in the context of ND polyphony for the same reasons shown there (ambiguity and variability of alignment). The melodic features say nothing about the alignment between voices, the chord features nothing about melodic movement, and it is impossible to reconcile the two.

This study will propose an alternate method for modelling the CANDR dataset: a two-layer graph where each element can have any number of links to any other element. The first layer (N or "node" layer) is a directed acyclic graph (DAG) per-voice used for temporal relationships, where elements are linked to their direct successor with a weight of 1. For example a note B following an A would link A to B with $N(1)$. Even notes that are known to be longer than notes in another voice are still linked to their direct successor with $N(1)$ as the alignment between voices is unknown.[3] The weight of the link says nothing about the duration of the note — as that, too, is ambiguous in ND notation — only that one element follows another. The second layer (S or "synchronous" layer) is an undirected graph indicating synchronicity between voices. Elements that are known to be synchronous are linked together with $S(1)$. Tokenisation is identical to a monophonic melody or text in that individual notes and other stave

[2] In the music analysis program *jSymbolic*, this can be mapped onto the "M" category features (presumably for "Melodic") versus the "C" category features (for "Chord(al)") <https://jmir.sourceforge.net/manuals/jSymbolic_manual/home.html>.

[3] Plus the often overlooked fact that a note cannot simply be reduced to a single event, or even a set of note-on, note-off events.

elements are considered tokens like before, but here they are placed in a graph rather than in a token stream. The musical notation is therefore modelled with very few assumptions made about its operation: all that is implied here is that some elements succeed other elements, and that some elements are synchronous with other elements. No assumptions are made about duration or alignment other than what is explicitly indicated.

6.2. Embedding

Just as "3.7.4. Embedding" discussed the utility of *Word2Vec* for creating embeddings for language in the context of natural language processing, there are embeddings available for graphs. Indeed, there has recently been a proliferation of methodologies for learning suitable embeddings for graphs, all at an early stage of research (Makarov et al 2021). These methodologies are particularly suitable for taking large and unknown graphs and condensing the whole graph or single nodes into a learned vector form. Graph neural networks (GNNs) are a particularly new set of methodologies in the field of machine learning, first developed only in 2016 (Kipf and Welling 2016), and with new improvements continually advancing their capabilities in tasks as diverse as social network analysis (Li et al 2023) and molecular property prediction (Jiang et al 2021).

Cutting-edge research in MIR has in recent years begun using GNNs for MIR tasks, and scholars have found that the flexibility of GNNs allows them to develop methodologies for a wide range of applications. Jeong et al (2019) used a GNN to generate piano performances, Baró et al (2022) applied GNNs to OMR tasks, Cosenza et al (2023) used GNNs to generate polyphonic music, Buisson et al (2024) used attention networks to improve the performance of GNNs on music structure analysis tasks (i.e. music sectionalisation and grouping), Lim et al (2024) generated symbolic music using GNNs from a dataset of pop music, and Karystinaios & Widmer (2024) have distilled their GNN methodology for analysing CWMN scores into an open-source software toolkit.

Of particular relevance to this study is Karystinaios and Widmer's (2024) state of the art methodology, where a CWMN score is preprocessed such that each note is linked to other notes by a number of different relationships, e.g. the notes beginning at the same time (onset), a note beginning while the other is sounding (during), one note follows another in time (follow), and a note beginning after a silence (silence) (Karystinaios and Widmer 2024). This is represented as a heterogeneous graph, which can then be passed into a GNN such as a graph convolutional network (GCN). Karystinaios & Widmer (2024) use neighbour sampling to increase efficiency by batching multiple small graphs together. This consists of extracting subgraphs from the larger graph structure each at a certain "hop" from the target node.

It would be ideal to use this GNN methodology directly on this dataset of ND polyphony — as the data structure used for the symbolic notation is already a graph structure — but it is not possible to use methodologies such as Karystinaios and Widmer's (2024) unaltered. This is because their methodology begins by converting encoded symbolic notation from a hierarchical tree structure such as MusicXML or MEI into their custom graph format, and this graph format replicates the same data structure relationships as the source files: "onset", "during", "follow", and "silence". Although we can conceive of the N and S layers in this study's format as analogues of Karystinaios and Widmer's (2024) "follow" and "onset" relationships respectively, the other two relationships in their model are not meaningful in the context of ND notation as modelled in this study. This could be worked around by converting the notation into a hierarchical format that could then be reconverted into this four-relationship graph format, but this would require making assumptions about the alignment between voices in the ND notation, and this was precisely what modelling the notation as a graph in the first instance was attempting to avoid.

This study, therefore, will design an altered and simplified methodology, based upon the work of Karystinaios and Widmer (2024), that more accurately reflects the data modelling decisions made for ND notation in "6.1. Modelling". Rather than modelling four different relationships for each element, this study will model the N and S layers as two different relationships within the notation graph. Each notation graph will be modelled as a set of elements — not just notes but also *divisiones*, clefs, and accidentals — linked to each other using the N and S layers previously described. Each notational element is a node in a graph, $e \in G$, where G is defined as an attributed heterogeneous directed graph, with feature vectors for each node. This feature vector is initialised to a one-hot encoded vector of the node type. One-hot encoding is a particular method of encoding categorical data as a binary vector, where a 1 is placed in an otherwise zero vector to indicate a particular category. An edge within the graph links two nodes, each with respective feature vectors, on either the N or S layer. As the S layer is undirected but the N layer is directed, S layer links will be reproduced as two identical links in opposite directions.

Whereas previous work also contains links for two notes sounding at the same time (during), this simplified methodology will only contain links for two notes beginning at the same time (onset). This will mean that there are far fewer links in the graph overall, and this would reduce the amount of message passing that can occur between nodes at each layer. Karystinaios and Widmer (2024) see success with a sampling hop value of 2 but, due to the fewer connections in this study's graphs, this value will be preliminarily increased to 5 so as to allow message passing (the core feature of GNNs) to propagate more quickly across sparser graphs such as those in this study. For example, a 2-hop

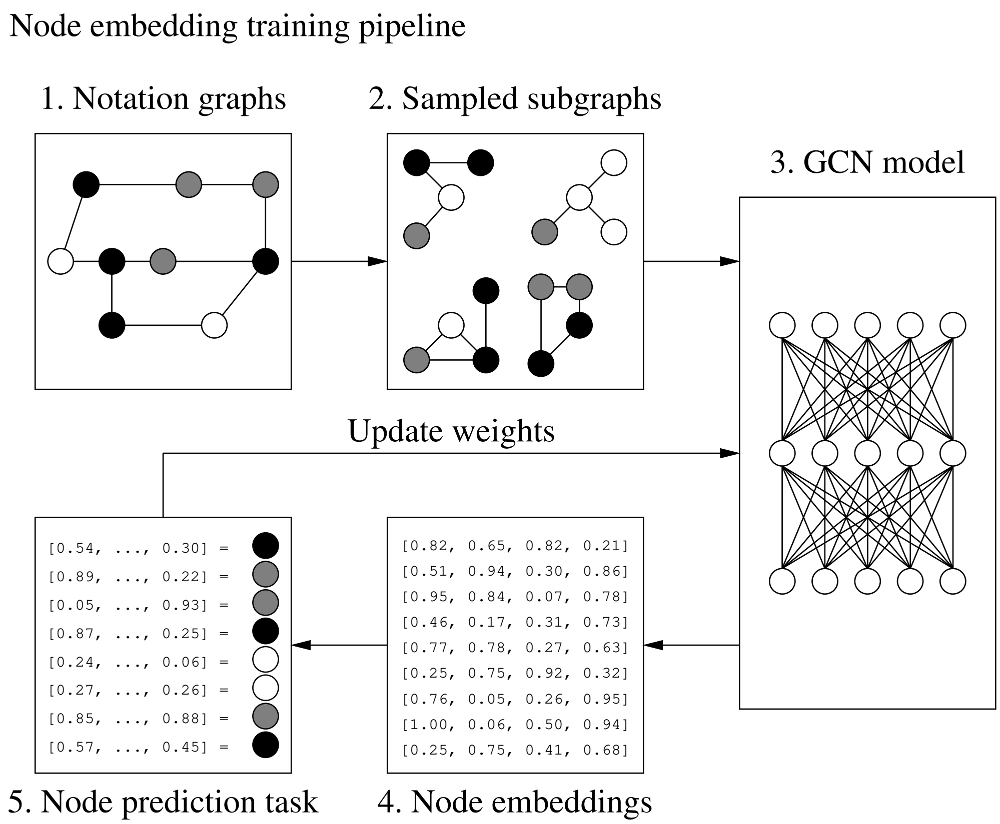

Fig. 6.1

would only allow the note next to a synchronicity to be given data about that synchronicity, but a 5-hop would allow notes four steps away to receive information about *S* layer links.

Three different autoencoder GNN models of various depths will be evaluated to generate node embeddings from these sampled subgraphs (see the training process in Fig. 6.1 and the model architectures in Fig. 6.2). Each model learns a node embedding of size *v*, i.e. a learned vector representation such that each node can be mapped into a fixed-length vector of size *v*. These node embeddings can subsequently be used as representations of notational elements in later stages of the methodology. In the broadest sense, these autoencoder models will take categorical information about notational elements and their surrounding context, and learn fixed-length vector representations that encode this same information in a condensed and fixed vector format. This will allow the categorical labels of the elements that follow a notational element (*N*) and the elements that are synchronous with an element (*S*) to be represented together as real-valued numbers which are more conducive for further machine learning than categorical features. The first part of each model learns an initial embedding of size *v* based on the node type, and this is concatenated with the node feature vectors. Each model then passes this vector through a series of GCN layers to

Three candidate GCN autoencoder models for node embedding training

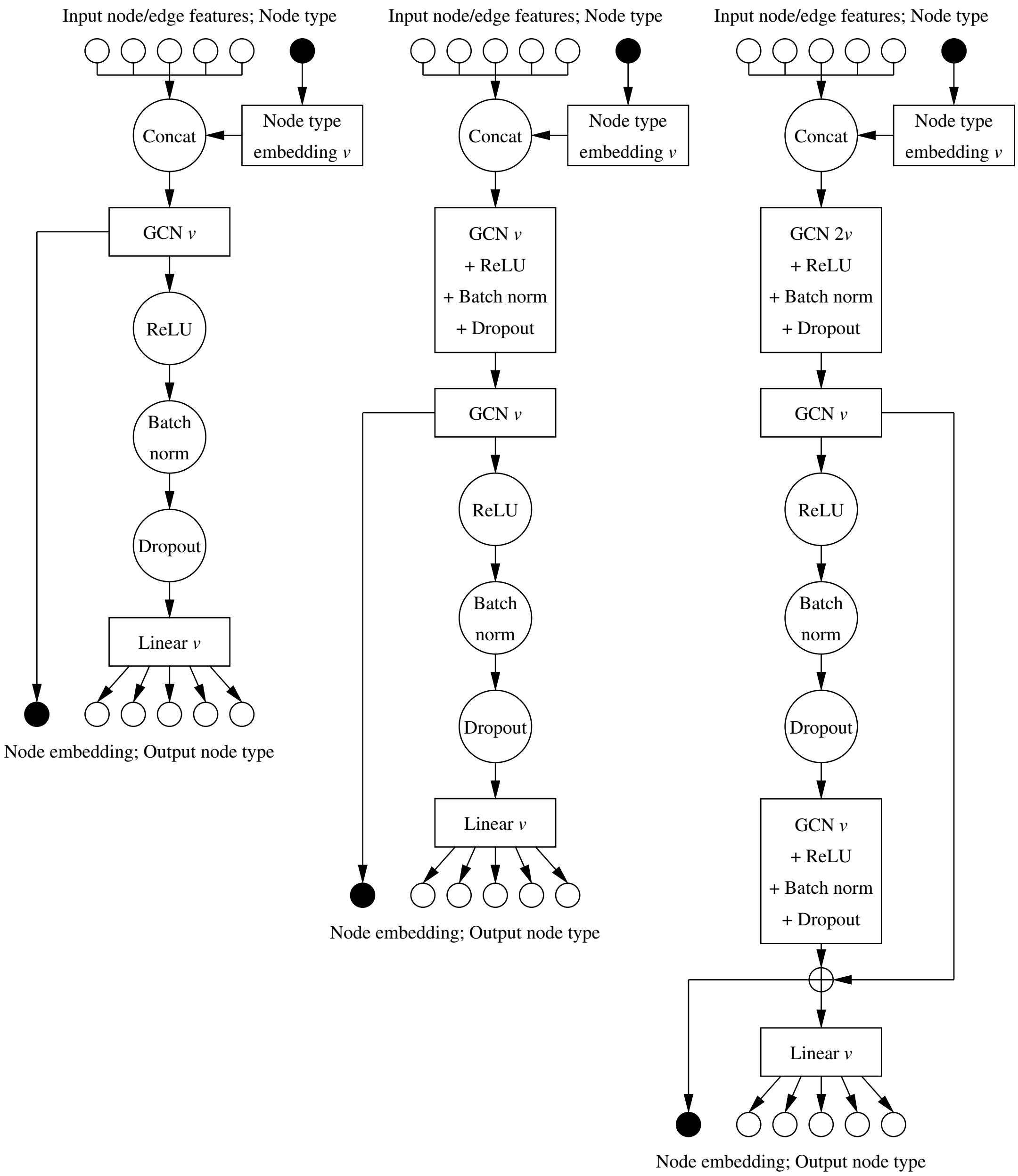

Fig. 6.2

encode information about the graph structure:

1. The simplest model uses a one-layer GCN of size v with ReLU activation, batch normalisation, and dropout.
2. The second model employs two GCN layers each of size v with ReLU, batch normalisation, and dropout.
3. The most complex model uses three GCN layers of sizes $2v \to v \to v$ with ReLU, batch normalisation, dropout, and a residual connection between the second and third layers.

Each model then decodes these embeddings using a single dense layer of size v into node types, and a cross entropy loss function is applied to the model's performance on a node prediction task. The best model for this study will be evaluated in "Chapter 7: Results and Analysis" by performing cross-validation, calculating clustering metrics, as well as downstream task performance in musical reuse detection.

6.3. Alignment

The outcome of this process will be a function that maps every notational element into a fixed-sized vector that represents the learned features created using the GNN model, much like natural language processing (NLP) methodologies such as *Word2Vec*. The advantage of using a GNN embedding here is that each vector will represent not only its own element type (such as A, B-flat, *divisione*, etc.), but also will encode features about the surrounding neighbourhood, and connections to other elements.

The node embedding methodology therefore will convert the representation of symbolic notation from a categorical data domain (node types and connections) into a continuous domain (fixed-size vectors). This will enable standard machine learning methodologies to be more readily applied to these vector representations. For example, similarity between two elements in the dataset will be calculated using a distance metric. The most common metric for comparing vector embeddings is cosine similarity,[4] and the similarity between any pair of vectors as created by the node embedding methodology can thus be calculated using this metric. It is important to state however, that a number of different metrics exist and, although cosine similarity is the standard for comparing vectors, it is not de facto the only possible or useful metric (Steck et al 2024). Nevertheless, cosine similarity is the natural choice for comparing the output vectors from the node embedding methodology, and although other metrics may provide different results, they will not be explored here.

The combined application of GNN node embeddings and cosine similarity will provide a methodology for comparing any two elements within the dataset. However, musical reuse is manifested in similarity over long passages, sometimes amounting to hundreds of notes, not just single elements. A methodology will therefore be required to find the most similar passages, namely, by comparing the similarities between elements using the node embedding representations. In "3.7.3. Sequence alignment" I outlined sequence alignment methodologies that minimise the edit distance between two sequences, such as the Needleman–Wunsch algorithm. However, these algorithms operate on categorical sequences of data (they were initially designed for comparing DNA sequences in bioinformatics) (Needleman and Wunsch 1970). As a result, comparing two dissimilar elements always results in the same score; a mismatch results

[4] See Žižka et al (2020, pp.239–240), a book which, likely not by chance, includes the words "cosine similarity" in large letters on its front cover.

Example of two similar, but misaligned time series, with dashed lines indicating DTW alignment

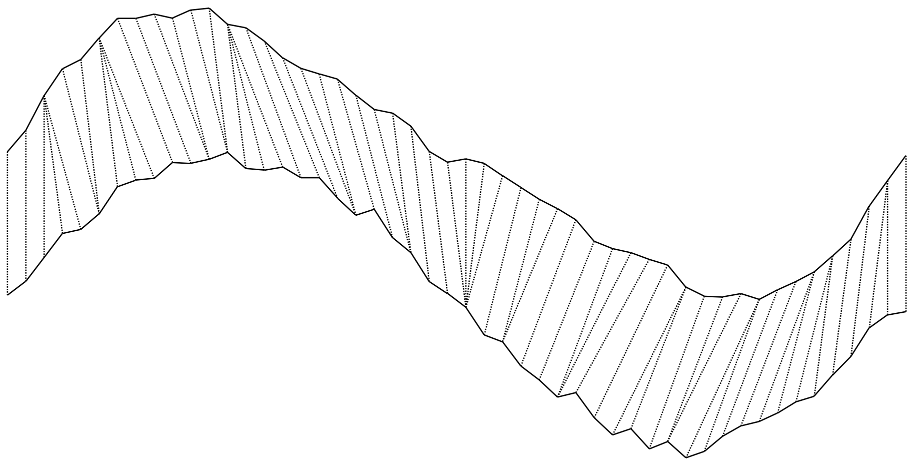

Fig. 6.3

in a fixed penalty (typically -1). Two similar but non-identical elements are scored the same as two elements which are entirely dissimilar. The output from the node embedding methodology is not this form of categorical data, but continuous vectors: none are identical but each have some level of similarity.

A cognate for sequence alignment in the case of vectors is Dynamic Time Warping (DTW) (Botsch 2023, pp.82–91). DTW is commonly used in time series analysis to align two sequences to each other using a cost function. Fig. 6.3 demonstrates this for a simple, univariate time series. Both are ostensibly sine waves, but each has some noise and they are offset by a nonlinear time distortion. Here, DTW is capable of creating a coupling between the two signals all the while keeping the order the same. It is also possible to apply DTW to multivariate signals such as vectors using a metric such as cosine similarity.

One possible alignment methodology, then, will be to flatten the notation graphs into a linear sequence and apply the node embedding methodology to each element. Each setting will therefore be represented by a multivariate time series of size $[n, v]$, where n is the number of nodes in a notation graph and v the size of the embedding vector. DTW will then be applied with cosine similarity to each pair of flattened sequences to generate best-fit alignments. Such a methodology has seen some use before, most recently in MIR to align note representations (Peter and Widmer 2024). However, this process works entirely on a single time series and, as mentioned, the polyphonic representation of symbolic music as encoded in the notation graphs must first be flattened. Again, the issues outlined in “3.7. Polyphonic problems” must be confronted: flatten how? and is this a suitable methodological approach when so much

Example of two similar distributions, with dashed lines indicating OT alignment

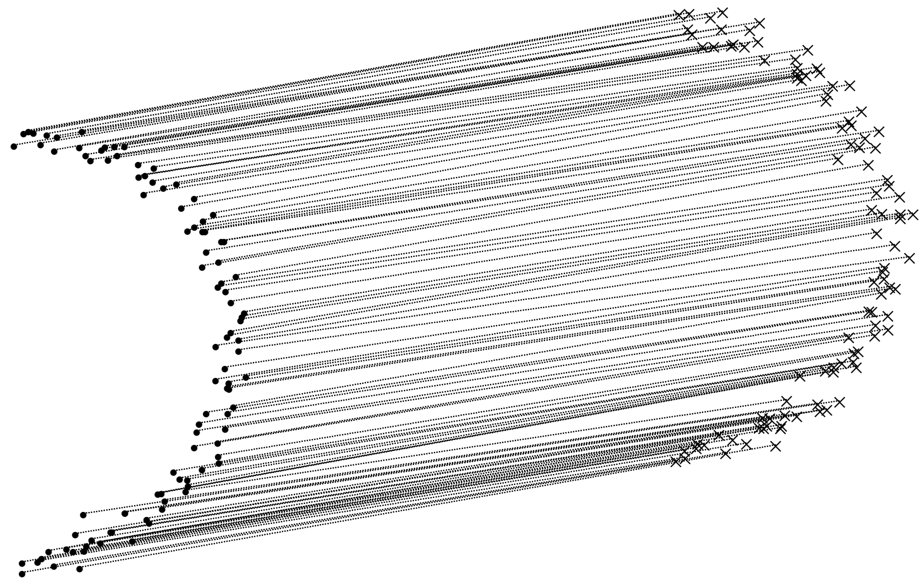

Fig. 6.4

of the previous work in this study has been done precisely to avoid representating ND notation using constrained data structures and convenient shortcuts? Although DTW is a possible alignment methodology, it is by no means the optimal solution.

An alternative methodology that does not have these issues can be found in a slightly related problem in transportation theory called Optimal Transport (OT). OT is concerned with calculating the most efficient way of moving one distribution to another with the minimal cost (Villani 2009, pp.29–37). There are many solutions to OT, but a common methodology is called Earth Mover's Distance (EMD). A cost matrix is calculated between all pairs of both distributions, and the best coupling between data points is found that minimises the overall cost.

Fig. 6.4 demonstrates OT for two simple 2D distributions. Both distributions are of the same half-moon shape, and OT can calculate the best one-to-one matching between the distributions. There are related algorithms for unbalanced OT (where the number of points in each distribution is not equal) and partial OT (where only some of the points should be matched) (Chizat 2016). The relevance of OT to this study is that the GNN embedding methodology created a fixed-size vector for each element in the notation graph, where similar elements are deemed to have a high cosine similarity. An alternate alignment methodology could therefore use OT to calculate the best matching between any two notation graphs by using the node embeddings to represent each notation graph as a distribution in a v-dimensional latent space. By plotting the vector embeddings of all the nodes in a pair of notation graphs, unbalanced OT with a cosine similarity metric would be able to generate an alignment between the two that minimises the distance between them.

The issue with this approach is that by plotting the elements as node embeddings, the relative location of each node within the graph would be lost, save what is encoded regarding the local neighbourhood from the node embedding process. A naïve OT on the node embeddings as distributions would ignore the overall temporal structure of the notation graphs, and match elements entirely out of sequence. The advantage of DTW on the other hand is that, although it falsely linearises the notation, none of the resulting matches cross.[5] OT has no such constraints on monotonicity and would thus match any two points together as long as they fit the overall transport plan. However, DTW's monotonic contraint can also be interpreted as a disadvantage, where out-of-order notes, ornamentations, repeats, and skips can confuse the process and generate misleading results (Peter and Widmer 2024).

This study will therefore extend the OT methodology by augmenting the learned embeddings with positional embeddings using two candidate methodologies. Both methodologies begin by assigning temporal markers to each node in each notation graph. Let $G_N = (V_N, E_N)$ represent the notation graph at layer N, where V_N is the set of nodes, and $E_N \subseteq V_N \times V_N$ the set of directed edges. The "start nodes" (A) are defined such that:

$$A = \{v \in V_N | \forall u \in V_N, (u, v) \notin E_N\}$$

and "end nodes" (B):

$$B = \{v \in V_N | \forall w \in V_N, (v, w) \notin E_N\}$$

Let $Q(a, b)$ be the set of all simple paths (that is, paths without repeated nodes) from $a \in A$ to $b \in B$. For a node $v \in V_N$, its position fraction is defined as the minimum normalised position of v across all paths it appears in. For a path $p = (v_1, v_2, ..., v_k)$, the position fraction of v is:

$$T(v) = \frac{i-1}{k-1} \text{ if } v = v_i \text{ in } p$$

where $T(v) \in [0, 1]$.

The first proposed positional embedding methodology can be described as a simple angular augmentation process, where the original embeddings $e(v)$ are augmented with two further features that represent the position fraction $T(v)$ as a fraction of a half

[5] Although there are lines that go to the same point in Fig. 6.3, none cross each other.

rotation. A scaling parameter (Φ) is defined as the power of this temporal augmentation. The optimal value of this parameter will be found experimentally in "Chapter 7: Results and Analysis". The two temporal features (λ) are defined as:

$$\lambda(v) = \sqrt{\Phi} \cdot \begin{bmatrix} \cos(T(v) \cdot \pi) \\ \sin(T(v) \cdot \pi) \end{bmatrix}$$

and the resulting augmented embedding:

$$e'(v) = \begin{bmatrix} e(v) \\ \lambda(v) \end{bmatrix}$$

The second proposed methodology is a more complex embedding commonly used in modern transformer architectures called Rotary Positional Encoding (RoPE) (Su et al 2024). RoPE has produced promising results in the NLP contexts of modern AI systems over absolute positional embeddings (Touvron et al 2023), and so may have similar improvements to offer in this case. Whereas the original seminal transformer paper that revolutionised post-2020 AI used absolute sinusoidal embeddings similar to the first methodology presented here (Vaswani et al 2017), RoPE uses a rotation matrix to transform the token embeddings in the embedding space by a position-dependent angle. RoPE can therefore be used to transform the GNN embeddings by adding positional encoding without the need for extra features as in the first methodology.

6.4. Linearisation and classification

Each of the three models described in "6.2. Embedding" will be evaluated in "Chapter 7: Results and Analysis" using both DTW and OT. It is known from recent work that DTW is capable of aligning note representations in music, building from foundational work in score following (Peter and Widmer 2024), but OT is entirely untested on symbolic music. DTW will therefore be used as a baseline methodology for this study, and OT evaluated against it. Furthermore, the two different positional embedding methodologies for OT will also be evaluated. Each of these positional embedding approaches have a parameter to optimise: the angular augmentation's scaling parameter (Φ) and, although RoPE commonly uses a base of 10,000 for its decay parameter (Su et al 2024):

$$\theta_i = 10000^{-2i/d}$$

more recent research is investigating how this parameter might be controlled for improved performance (Liu et al 2024). It may therefore be possible to optimise the results by modifying this base. The various methodological permutations and parameter options will be selected here by evaluating each possibility using a grid search, and selecting the optimal parameters via their score on a downstream task generating separate distributions both for the CANDR dataset as a whole and a known subset of pairs of similar transmissions.

The resulting dataset will comprise a set of cosine similarities between pairs of aligned notation graphs within the CANDR dataset. Passages of extended low cosine distance will indicate sustained alignments between the two graphs analysed, and therefore evidence for musical similarity. For a single pair of notation graphs, this series can be created by sorting the matches between a pair of graphs (*X* and *Y*) by the position fraction *T(v)*, and listing the cosine distances between each successive match. Let V_X and V_Y represent the sets of nodes for graphs *X* and *Y* respectively. The union of the nodes is $V = V_X \cup V_Y$. Define the sorted list *S* as:

$$S = \text{sorted}(V, \text{key} = T(v)) = \{v_1, v_2, ..., v_n\}, T(v_1) \leq T(v_2) \leq ... \leq T(v_n)$$

$M \subseteq V_X \times V_Y$ is the set of matched edges created from DTW alignment or OT, where $(u, v) \in M$ indicates that node u in *X* is matched to node v in *Y*. The result of the positional encoding (either by temporal augmentation or RoPE) is a function E that maps each node $v \in V$ to an embedding vector. For embeddings $E(u)$ and $E(v)$, their cosine distance is defined as:

$$\text{cosine\_distance}(u, v) = 1 - \frac{E(u) \cdot E(v)}{\|E(u)\| \, \|E(v)\|}$$

and this results in a list of cosine distances throughout the graph, ordered by positional encoding.

It will be possible then to plot similarity over time for each pair of notation graphs in the CANDR dataset. However, with a dataset comprising nearly 1,000 settings, the number of pairwise graphs created follows the binomial coefficient:

$$\binom{n}{2} = \frac{n(n-1)}{2}$$

For $n = 1000$, this would result in nearly half a million plots (499,500). Combing through all of these results to find those passages where cosine distance is low will not be feasible, and so a further process will be used to detect these passages automatically.

The first question to be asked is: what do we mean by a "low" cosine distance? As the bounds of cosine similarity are $[-1, 1]$ and cosine distance is defined as $1 - \text{cosine}$ similarity, cosine distance is bounded by $[0, 2]$. However, what may be construed as a low distance in the context of one comparison — and therefore evidence of musical

Hypothetical criterion for musical similarity

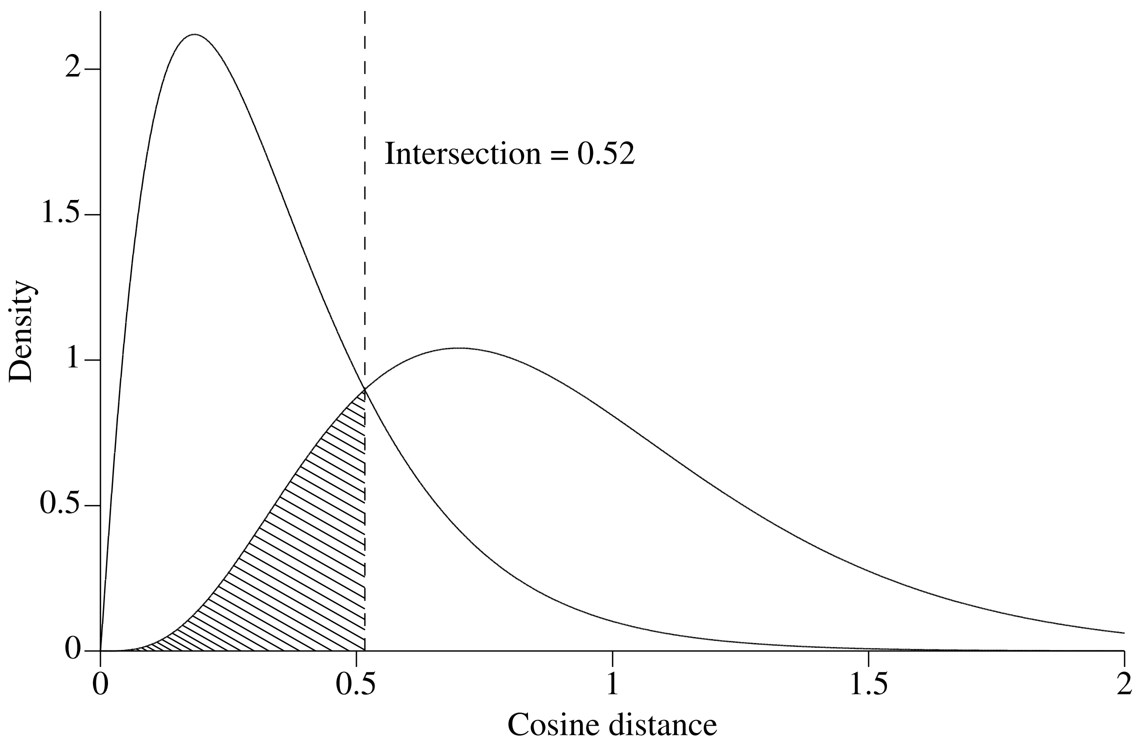

Fig. 6.5

reuse — may be interpreted as a higher distance in a different comparison. It may be possible to use the dataset itself to find out where this cutoff criterion for further study should be placed, by comparing the distribution of all pairwise comparisons (assumed to be mostly unrelated) to a small selection of pairs that are known to be related (such as by clausula substitution or indirect concordances). If the cosine distance between known similar pairs (J) is distributed differently to all pairs (K), then a threshold cosine distance can be decided upon; all cosine distances lower than this point can be classified as similar, and all distances above as different. This point could be calculated as the point at which the likelihoods of the two distributions are equal (see Fig. 6.5):

$$P(x|J) = P(x|K)$$

The issue with this simple thresholding process is that it is far too reminiscent of the all too inflexible criteria critiqued in “2.9. Melodic formulae”: simply deciding on a threshold criterion for what constitutes “similarity” vs “difference” does not consider the surrounding context, and whether this similarity is significant within the setting as a whole. For example, in Fig. 6.5 a not insignificant proportion of the flatter distribution is below the intersection (shaded area). Despite its density being less than the narrower distribution, the probability that a sample has been drawn from this distribution is far from zero. We cannot tell from a single sample alone whether it is evidence of musical

Example Hidden Markov Model

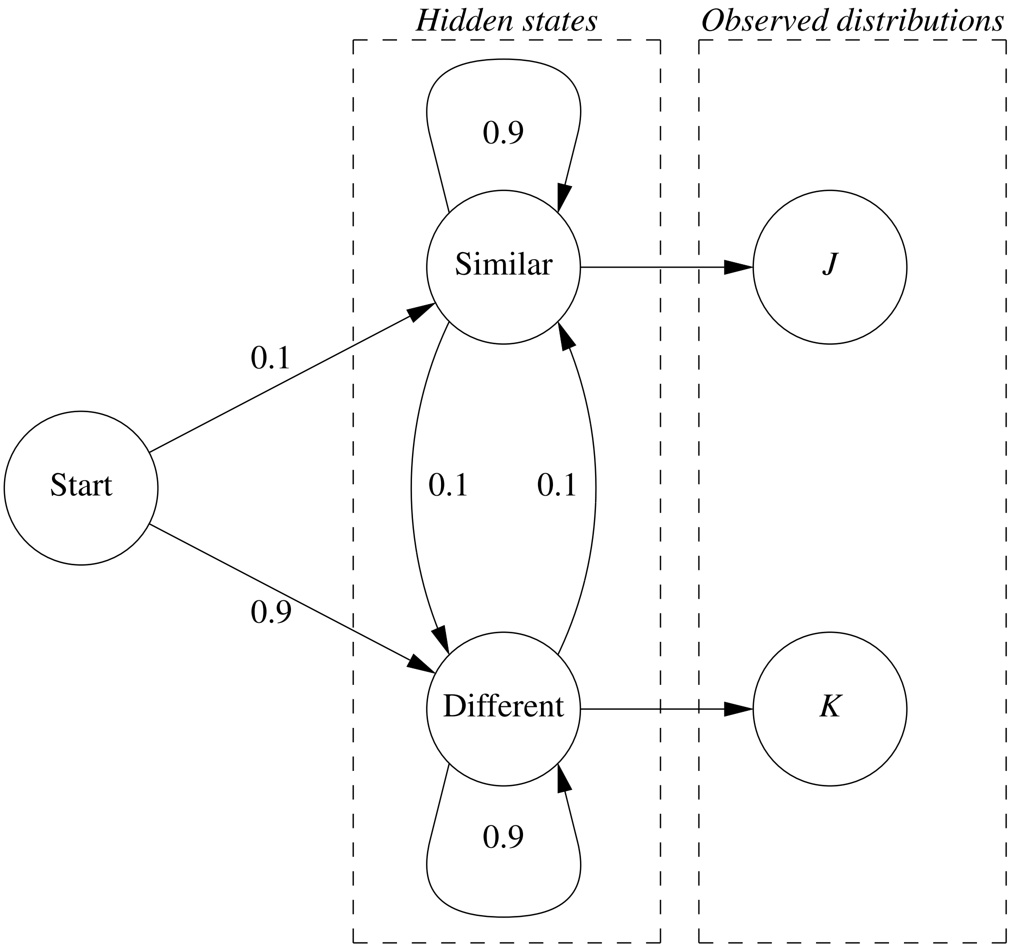

Fig. 6.6

reuse. The alternative approach that will be taken here will be to take the surrounding context into account by modelling the hypothetical distributions J and K as two states within a Hidden Markov Model (HMM).

Fig. 6.6 demonstrates such a model applied to the idea of similarity vs difference. There are two observed distributions, J and K, representing distributions of "similar" and "different" cosine distance respectively. These are produced by two hidden states: J is produced by one hidden state ("similar"), and K by another ("different"). The list of cosine distances extracted from the alignment process can be conceived of as a noisy signal that switches between these two states: it is usually producing samples from the K distribution (indicating a "different" hidden state), but on occasion may produce evidence of musical reuse in samples from the J distribution (indicating a "similar" hidden state). However, these distributions have some overlap, and we do not know at any one point which distribution the signal is being sampled from. We want to calculate the most likely sequence of states that produced an input signal.

A categorical probability distribution defines the probability of beginning in any one state — given here by arrows from the circle labelled “Start” — and a transition matrix defines the probability that a succeeding sample in the signal will come from the same distribution or the other state. In Fig. 6.6, the HMM is configured to be particularly “sticky”, that is, the transition matrix prefers to stay in the “different” state, with only 10% probability of transitioning to a “similar” state. Then, when it is in the “similar” state, it has only a 10% probability of transitioning back to the “different” state. This highly weighted transition matrix ensures that the HMM must receive a strong signal to the contrary in order to change state, so that it is not constantly flip-flopping between states and predicts long, contiguous passages as being in a single state.

The probability distributions of J and K for the HMM can be decided upon by observing the empirical distribution of these two datasets. These distributions can then be used to train the HMM in a semi-supervised manner, by assigning categorical weakly informative priors to the samples: for example, a sample from K can be given a prior of [0.1, 0.9], indicating our prior belief that this sample most likely came from the “different” state, but it is not impossible that it was emitted by the other “similar” state. Values of the exact priors assigned, as well as the transition matrix of the HMM, will be found through experimentation.

6.5. Programming toolkit

The CANDR analysis toolkit is written in *Python* and not only implements the above methodology more robustly as a set of classes and library functions, but also provides a programming framework for interacting with the methodology in an exploratory, extensible, efficient, and user-friendly manner. With several options for importing and exporting data from the toolkit, the library allows users to seamlessly integrate the methodology into their pre-existing analysis workflows. The toolkit is designed with exploratory analysis in mind, and therefore provides an intuitive programming framework that facilitates the interactive exploration of data. Users can easily manipulate their datasets into forms conducive to analysis within the methodology, and then obtain valuable insights through an iterative and dynamic process.

Due to the extremely large volume of data processed as part of the methodology, the CANDR toolkit is designed to handle large datasets and optimise computational resources. The framework leverages several packages from *Python*’s rich ecosystem of scientific computing libraries to streamline the analysis process and provide results that are interoperable with other systems, enabling researchers to derive meaningful results efficiently.

Moreover, one of the reasons that the framework was implemented as a *Python* module is to promote extensibility, allowing users to customise and extend its functionalities to suit their specific needs. By providing a modular structure, the toolkit enables the incorporation of further algorithms and statistical techniques, as well as visualisations. This empowers users to adapt the methodology to further domains and research questions. This study asks one particular and narrow question about this repertory that this programming toolkit can help to answer, but the aim of creating the programming toolkit is also to provide a framework for further study. For example, an improved methodology for studying this same research question may be developed, and the toolkit is designed in such a way as to facilitate the substitution of methodologies by structuring the process as a series of interoperable components.

The toolkit has also had to create solutions for working with such a large dataset, and all manner of other research questions could be formulated as components of this toolkit. For example a study of harmony in this corpus could very easily create a program for extracting statistics relating to intervals in the dataset without first having to deal with the issues of size and scale.

Finally, the CANDR dataset is not the only corpus that could be analysed using this methodology: other ambiguous or early polyphony, or even CWMN could be processed this way, perhaps by re-integrating the "during" and "silence" relationships from Karystinaios and Widmer (2024). Therefore, the modules that fetch data from the CANDR API are separated in the toolkit into individual modules so that they can be substituted for other data sources. For example, this methodology may be suitable for comparing the complex and nuanced problem of musical reuse within the music of Charles Ives (Burkholder 1995), and so if a new data source of Ives' music was created, it could very easily be added to support this analysis.

6.6. Pipeline and PipeObject

The core of the programming toolkit is a lightweight custom pipeline framework designed for creating and executing data analysis stages, as this allows various different components within the toolkit to be easily connected together both in sequence and in parallel, affording both the exploratory extensibility and interoperability of the system. This pipeline framework was created as there are often multiple stages to the analysis, such as pre-processing and parsing, as well as several stages of independent analysis. However, most pipeline frameworks currently in existence are focused on task delegation and processing, either at a local level by parallelising array processing,[6] or at a higher level by delegating batches of data to nodes within a High Performance Computing (HPC) cluster.[7]

The focus of this pipeline framework however, is on efficiently streaming data

[6] Such as *Joblib* <https://joblib.readthedocs.io/en/latest/parallel.html>.
[7] <https://docs.dask.org/en/stable/deploying-hpc.html>.

between tasks within a local machine. For example, for this project's analysis, hundreds of settings of music must be downloaded from the CANDR website server, parsed, pre-processed, and analysed through a complex interplay of processes. Pipeline or workflow frameworks typically accomplish this by distributing the entire set of tasks to their worker nodes, waiting for the results, and collecting the results together.[8] Only once the results are collected can the next stage be processed. The CANDR pipeline is instead focused on a different dimension: that of efficiently streaming results from one stage to the next as soon as possible.[9]

The CANDR pipeline uses *Python* generator functions, each within its own process, to iterate through input and generate results as soon as they are ready. The pipeline provides functionality for connecting various components of the pipeline using shared memory queues, allowing the results from one process to be instantly sent to the input of another. However, this has trade-offs in the fact that large amounts of data have to be copied between processes adding a sizeable memory overhead to each process. A single process can do any amount of processing on its data without having to copy, but passing the data between processes necessitates copies being made, one for each process. As more data is fed into the system, it becomes difficult for *Python*'s memory management system (the garbage collector) to keep up, and the system runs the risk of running out of working memory. This is a very typical example of a space–time trade-off, where a program may be run more quickly by using more memory. This has been mitigated somewhat by carefully invoking the garbage collector manually within long-running processes, but remains an issue for the toolkit, where processes that transfer a lot of data between *PipeObject*s are prone to running out of memory.

At the centre is a *Pipeline* class, which orchestrates and connects components together. Where necessary, it uses the *Python* multiprocessing library to achieve parallel computing.[10] The *Pipeline* class' key functions are `add()` for adding components to the end of the pipeline and `execute()` for running the pipeline from beginning to end. `add()` creates a shared memory queue between the output of the current end of the pipeline and the input of the new component such that results from one component flow directly into the input of the next. `execute()` starts all of the components in individual processes — in reverse order to prevent congestion — and then waits for them all to finish. At the same time, the *Pipeline* receives updates from the components and displays these updates as a table to the user.

There is then an abstract class *PipeObject* which forms the basis of all components to be added to a *Pipeline*. *PipeObject* contains the utility functions for interfacing with

[8] See the outline provided for *Dask* <https://distributed.dask.org/en/stable/journey.html>.
[9] Since creating the CANDR pipeline, I have found a more comprehensive and general solution to this problem already exists in *Apache Flink*. See <https://flink.apache.org/>.
[10] <https://docs.python.org/3/library/multiprocessing.html>.

the shared queues and the *Pipeline* orchestrator, as well as common functions shared between components such as `set_status()`. However, *PipeObject* is designed to be sub-classed into concrete classes that perform meaningful work in the pipeline. For example, here is a simple derived class that adds a constant number to its input:

```
class AddN(PipeObject):
    """A PipeObject that simply adds n to what is input"""
    def __init__(self, n, **kwargs):
        self._n = n
        super().__init__(**kwargs)
    def process(self, input_data: int) -> Iterator[int]:
        yield (input_data + self._n)
```

At its initialisation, it takes in a number, *n*, which will be the constant to be added. It then defines a function `process()` which takes in `input_data` from the input queue — administered by the abstract *PipeObject* — computes a result which is `input_data + n`, and then "yield"s the result, i.e. returns the result as a generator. From this basic example, it is easy to imagine a more complex analysis component that generates numerous statistics at different times, even for one input value. Rather than saving all those statistics within the function until the analysis is complete, they can be elegantly streamed to the next stage of the process.

6.7. Key components

In "4.6. An overview of CANDR", I showcased some example *Python* scripts that use the programming toolkit to perform meaningful analysis on the CANDR dataset. Each script first sets up a pipeline and progress updater like so:

```
from candr_analysis.utilities import pipeline, progress

prog = progress.CLIProgress()
processes = pipeline.Pipeline(progress=prog)
```

Components can then be added to the pipeline using `processes.add()`, and the pipeline run using `processes.execute()`. The first script used three components: a) `candr_analysis.scrapers.candr.rdf.setting.Setting`, b) `candr_analysis.transcribers.candr.MEI.Setting`, and c) `candr_analysis.transformers.pickle.PickleSaver`. The second script introduced d) `candr_analysis.transformers.pickle.PickleLoader`, e) `candr_analysis.utilities.pipeline.Multiplexer`, and f) `candr_analysis.parsers.MEI.MEIParser`. The third script used g) `candr_analysis.transformers.MEI.ego_graph` and h) `candr_analysis.transformers.MEI.ego_graph.EgoGraphsToData`. The final script introduced i) `candr_analysis.analysis.graph.graph_network.GraphRGCNNodePred` and j) `candr_analysis.graph.graphstatistics.GraphStatisticSaver`. The following sub-sections will describe the function of each of these components within the methodology.

6.7.1. Retrieving RDF and MEI

One of the ways in which data can be imported into the programming toolkit is by fetching data directly from the CANDR website server. The CANDR website pages are marked up using RDFa (Resource Description Framework in Attributes) to provide rich metadata for programmatic access to the website (Sporny et al 2015). In the first script, *rdf.Setting* (a) and *MEI.Setting* (b) use these features to download data directly from the website server. The setting scraper navigates directly to the setting list page at `/browse/setting/` and parses the RDF data contained therein. It extracts the RDF triples for each setting and creates individual RDF graphs for each setting.

The MEI transcriber parses an individual RDF graph and finds from this the URI of its MEI transcription endpoint. It then fetches the MEI data directly from the server and creates an internal *MEITranscriptionRecord*, which is a simple wrapper around the unparsed MEI markup.

6.7.2. Saving to and loading from disk

All the elements of the pipeline can be linked and executed together, but to save time for repeated runs and for debugging and inspection of intermediate results, it is possible and expedient to save and load results to *Python pickle* files. These are serialisations of *Python* objects and can be loaded to and from memory at will.[11] Therefore, the first two scripts demonstrated in "4.6. An overview of CANDR" use *PickleSaver* (c), a *PipeObject* that takes its input and saves the intermediate result to disk as a *pickle* file. The last two scripts use the inverse: *PickleLoader* (d) which takes a *pickle* file and loads results into a pipeline.

PickleSaver is relatively simple: on receiving an input, it calls `pickle.dump()` on that input to serialise it directly to file. The file is appended so that any number of serialisations can be stored in one *pickle* file. *PickleLoader* achieves the opposite by first validating the input file before processing. It does this by attempting to de-serialise all the records in the file one-by-one, throwing an error if the file is corrupt. This is also useful for showing progress, so that the pipeline knows how many records to expect. It then rewinds the file to the beginning and de-serialises the records again to its output, sending a sentinel value at the end to indicate the end of the file.

[11] <https://docs.python.org/3/library/pickle.html>.

6.7.3. Parsing and multiplexing

The *MEIParser* (f) ingests the *MEITranscriptionRecord*s and generates *MEIParsedRecord*s, being complete graphs of the *N* and *S* layers of the data. It first parses the XML using *Python*'s *ElementTree* class and passes this to an *MEIValidator* class.[12] This validator class ensures that the MEI file is in a format that the parser can understand, and that each of the elements has the necessary attributes. It does this by using custom functions for validating each kind of element, such as `_validate_clef()` and `_validate_note()` containing numerous "assert" directives stating facts about the attributes that should be true, such as the `@pname` of a note being one of "a", "b", "c", "d", "e", "f", or "g".

The *MEIParser* then uses the library *NetworkX* to create a multi-digraph.[13] This is a directed graph in multiple layers. All the *N* links are stored on one layer, and the *S* links on another. The *S* links should be undirected however, but *NetworkX* has no construct for some layers being directed and others undirected. This is worked around by creating two links for each undirected link: one in one direction, and one in the opposite. The *MEIParser* iterates through the MEI document and inserts sequential elements into the *N* layer. For *S* links it maintains a dictionary of XML IDs: on parsing a node, the dictionary is checked to see if that XML ID should be linked to other nodes, then any `@synch` nodes are either looked up in the dictionary to be linked, or added for later linking (assuming that they will be found later in the document). The graph is copied into an *MEIParsedRecord* and output to the next stage.

The *MEIParser* is a compute-intensive task but is CPU-bound. There is no way in which this task can be easily parallelised per-task. The CANDR pipeline provides a *Multiplexer* component (e) for more traditional task distribution among multiple workers by providing an API translation layer for child tasks. Tasks can be added to the *Multiplexer* in exactly the same way as the *Pipeline*, and to the child tasks the *Multiplexer* provides an identical API. However, the *Multiplexer* maintains its own set of output and status queues and keeps track of the order of jobs. The *Multiplexer* retains buffers for the inputs and outputs for each worker so that if a job finishes early the worker is not left idle but can receive a new job. The result is then assembled back into order for the *Multiplexer*'s output queue. To the *Pipeline* then, the *Multiplexer* acts just like any other *PipeObject*. In the second example script (p.154), an *MEIParser* is instantiated for every CPU in the system, so that the parsing process is parallelised and uses all available CPU cores. On the machine that was used to complete the analysis of the CANDR dataset therefore, it increased the processing speed of the parsing step eightfold.

---

[12] <https://docs.python.org/3/library/xml.etree.elementtree.html>.
[13] <https://networkx.org/documentation/stable/reference/classes/multidigraph.html>.

6.7.4. Embedding

These parsed MEI notation graphs are then passed to a sampling process contained within the *EgoGraph* (g) component. This simply uses *NetworkX*'s `ego_graph()` function to generate ego graphs for every node within a notation graph at a given radius.[14] However, these are not in the right format for using as training data for a GNN. *EgoGraphsToData* (h) therefore converts the dict-of-dict format preferred by *NetworkX* into the set of matrices required for a GNN: edge indices, edge types, and edge weights. The node types themselves are encoded as a one-hot vector, and these are given as initial node features to nodes in the notation graph.

GraphRGCNNodePred (i) ingests this GNN-ready data and computes class weights for the cross entropy loss using *scikit-learn*.[15] An 80:20 training:validation split is made in the data by random sampling. The models themselves are created using *PyTorch Geometric*, an extension of the popular *PyTorch* machine learning library to GNNs.[16] The models are transferred to a GPU (Graphics Processing Unit) for faster training using NVIDIA's CUDA (Compute Unified Device Architecture).[17] The models are trained over a number of epochs, and the cross entropy loss for each training batch is logged.

The models are then placed in validation, and the validation dataset is tested using the models. Two metrics are computed: a multiclass AUROC (area under the receiver operating characteristic curve) in a one-vs-rest fashion, and the Calinski–Harabasz index.[18] The AUROC is calculated on the node type output of the models, and informs us as to the autoencoder's performance on classifying node types from the validation dataset. The Calinski–Harabasz index is computed over the true node types and output node embeddings, and provides a relativistic impression of the clustering performance of the embeddings; better clustering will result in a higher index. Finally, all the nodes are run through the model to generate embeddings, and a number of statistics are output:

- Class weights for each node type;
- Cross entropy loss over the training process;
- Receiver operating characteristic (ROC) curve;
- Overall AUROC score;
- Calinski–Harabasz index; and
- Node embeddings for every node in the dataset.

---

[14] <https://networkx.org/documentation/stable/reference/generated/networkx.generators.ego.ego_graph.html>.
[15] <https://scikit-learn.org/1.6/modules/generated/sklearn.utils.class_weight.compute_class_weight.html>.
[16] <https://pytorch-geometric.readthedocs.io/en/latest/>.
[17] <https://pytorch.org/docs/stable/notes/cuda.html>.
[18] <https://pytorch.org/torcheval/stable/generated/torcheval.metrics.MulticlassAUROC.html> and <https://scikit-learn.org/stable/modules/generated/sklearn.metrics.calinski_harabasz_score.html>.

The *GraphStatisticSaver* (j) component receives all of these statistics and saves each to a log folder in CSV format.

6.8. Alignment and result extraction

The remainder of the methodology is a little more static, as it operates on datasets that can be contained within working memory, and therefore the analysis code is written in pure *Python* without the *Pipeline* class. Recall from "6.2. Embedding" that there are three node embedding models of increasing complexity to be evaluated. Each creates a set of node embeddings as a CSV file. In "6.3. Alignment", I demonstrated two different possible alignment methodologies — DTW and OT — and two different positional embedding strategies for OT — temporal augmentation and RoPE — each with a parameter (Φ and RoPE base). The optimal parameter selection is unknown, and so each combination of parameters must be evaluated, and the highest-performing result will be selected for further analysis in "Chapter 7: Results and Analysis".

As mentioned in "6.4. Linearisation and classification", the alignment process creates hundreds of thousands of sequences of cosine distances between pairs of settings. For most, we can assume that these cosine distances represent no relationship between the pairs. However, this study will take a small selection that are known to be related and compare the distributions between the known similarities (J) and likely differences (K). We want to select parameters that maximise the difference between distributions J and K. This difference can be quantified using metrics such as the Kullback–Leibler divergence (Kullback 2006), which measures how much information is lost when one distribution is used to approximate another, the Wasserstein distance, a special case of EMD that measures the minimum cost of transporting one distribution to another, and a simple measure of intersection:

$$\text{Intersection} = \sum_{i=0}^{n} \min(J_i, K_i)$$

that quantifies the amount of overlap between the two distributions. These metrics, plus the AUROC and Calinski–Harabasz scores from the node embedding process, will be used to select the best parameters for the study when iterated using a grid search. A *Python* script enumerates all the possibilities, such as model 1 with DTW, or model 3 with OT RoPE embeddings (base 1,000), and creates histograms for both the known similarities and dataset as a whole. The histograms are normalised so that although the sizes of both sets (all alignments vs know similarities) are orders of magnitude apart, they can be compared directly.

The HMM is trained using *Pomegranate*, a probabilistic modelling library for *Python* (Schreiber 2018). Typical HMM libraries such as *hmmlearn* contain a limited selection of available models for continuous emissions, such as Gaussian and Poisson distributions.[19] *Pomegranate*, on the other hand, allows any combination of distributions to be used in the HMM, even in mixture models, and training can be accelerated using CUDA.[20] A set of continuous distributions are selected from *Pomegranate*'s API: exponential, gamma, normal, and Poisson. They are fitted to both *J* and *K*, and the log likelihood is calculated for each as a measure of how well the distributions fit the data. A higher log likelihood score equates to a better fit to the empirical distribution. If none fit the data very well, then a series of mixture models can be created along the same lines.

Finally, *Pomegranate* is used to create an HMM with "sticky" transition probabilities, as outlined in "6.4. Linearisation and classification". The best-fitting distributions are selected and used as the emissions for the two states "similar" and "different". Data from *J* are given a prior privileging a "similar" state, and data from *K* a prior privileging a "different" state. The model is then trained using CUDA. The results are postprocessed using a simple custom function that removes some false positives, by filtering out any "similar" predictions where the length of the contiguous interval is less than some minimum value. Increasing this minimum value will reduce the number of results generated, and decreasing this minimum value to 1 will produce all results, regardless of how long the "similar" state is predicted.

6.9. Conclusion

This chapter has presented the second part of the methodology of this project, the original data analysis methodologies that have been used to process the CANDR dataset, and described its implementation in detail. With this available, it is therefore now possible to perform the requisite high-level analysis on ND polyphony and generate meaningful results to detect patterns of musical reuse. The implementation of this novel methodology allows for this analysis technique to be performed on the whole CANDR dataset as a corpus, and the evaluation of this methodology, analysis of the CANDR dataset, and creation of useful results will be discussed in the next chapter. This methodology, through its automated comparative analysis between settings, will therefore generate results that will help to answer the research questions and will provide evidence for a flexible conception of musical reuse within the ND repertory. After the techniques discussed in this chapter, this will be done through the extraction of broad patterns common between settings, and the changing expression of patterns from one setting to another.

---

[19] <https://hmmlearn.readthedocs.io/en/latest/tutorial.html#available-models>.

[20] <https://pomegranate.readthedocs.io/en/latest/api.html#pomegranate.gmm.GeneralMixtureModel> and <https://pomegranate.readthedocs.io/en/latest/faq.html>.

Chapter 7
Results and Analysis

With all the tools and data required at my disposal, the analysis of musical reuse patterns within the ND repertory can be performed. The task at hand, to find evidence of musical reuse beyond the typical clausula-concept substitution, has not only necessitated the creation of new datasets and bespoke tooling to facilitate the input and analysis of the notation, but also a reappraisal of the way in which the ND repertory has been studied in the past. This has included its place in historical narratives, its conceptual dissimilarity to CWMN, and the thorny issue of musical reuse that this study is concerned with. These tools will come together in this chapter to provide evidence for a new understanding of the breadth and depth of musical reuse within the ND repertory beyond what is already apparent from close reading of the manuscript sources. Therefore, this chapter documents the way in which I have used the tools created as part of CANDR to answer the research questions, presents the results generated, and analyses how these results provide evidence for further musical reuse within the ND repertory. This chapter moves away from the heavy methodological focus of the preceding chapters back to the core issue of this study, which is musicological, albeit viewed through an empirical and data-driven lens.

Using these tools, there are four stages that must be completed in sequence to create such an analysis, and this chapter will begin by documenting how this was achieved. First is facsimile image acquisition, being the fetching of the digital images of the manuscripts from their respective repositories on the host libraries' websites. Second is image importing, tagging, cataloguing and uploading these images into the CANDR system so that they can be processed. Third is data input, using a combination of manual input ("keyboarding") and the OMR system to trace the locations of notational elements on the images. Finally is the process of creating an analysis program using the programming toolkit, linking it to the CANDR database, running the program, and collecting the results.

7.1. Image acquisition

The initial aim of the study was to input and fully trace the three central sources of the ND repertory — **F**, $\mathbf{W_1}$, and $\mathbf{W_2}$ — and so it was these image sets that had to be procured. Many high-resolution facsimiles of medieval music are hosted via the *Digital Image Archive of Medieval Music* (DIAMM), which collates facsimiles, inventories, source descriptions from the *Répertoire International des Sources Musicales* (RISM), as well as bibliographies.[1] Unfortunately, no facsimiles of **F**, $\mathbf{W}_1$, or $\mathbf{W}_2$ are hosted there. However, both $\mathbf{W}_1$ and $\mathbf{W}_2$ are held in the same institution, the Herzog August Bibliothek, Wolfenbüttel (HAB), and DIAMM links directly to where images of these sources are hosted on the HAB website.[2] The HAB interface contains a link in the top right that links to a usage notice, which states that 'the digitised images of HAB holdings are published as public domain'.[3] It is therefore possible to use these images in CANDR without restriction. Nevertheless, I did provide attribution in the system for integrity.

A link is also provided on the HAB website to download sources' images as a PDF. The entire source can be downloaded this way. The HAB viewer contains three zoom modes: a zoomed-out minimum view, a default medium view, and a zoomed-in high resolution view. For the tracing process, I wished to obtain the maximum resolution, as this would provide more detail for accurate OMR and provide a clearer view for manual input. However, the PDF download contained the images only at their default medium view and not at the high resolution view. In the export modal, there is no way of altering the resolution of the PDF. However, viewing a facsimile at different zoom modes and inspecting the image source loaded showed that each zoom mode is saved as a different image on the server. For example, the first image is by default saved at <https://diglib.hab.de/mss/628-helmst/00001.jpg> but also has a minimum view at <https://diglib.hab.de/mss/628-helmst/min/00001.jpg> and a maximum view at <https://diglib.hab.de/mss/628-helmst/max/00001.jpg>. Each successive image is saved as a five-digit monotonically increasing integer. For $\mathbf{W}_1$, this begins at "00001.jpg" and ends at "00418.jpg". Each maximum resolution image could be saved individually, but downloading each one would be laborious. I streamlined this process by using a simple shell script:

```
#!/bin/bash
# Set the base URI
url="https://diglib.hab.de/mss/628-helmst/max/"
# Iterate a five digit integer from 00001-00418
for i in $(seq -f "%05g" 1 418); do
    # Download the JPG
    wget "${url}${i}.jpg"
done
```

This script downloaded all the images of $\mathbf{W}_1$ into a single local directory. The same was done for $\mathbf{W}_2$, but with the URI <https://diglib.hab.de/mss/1099-helmst/max/> and the

[1] <https://www.diamm.ac.uk/>.
[2] $\mathbf{W}_1$ can be found at <https://diglib.hab.de/wdb.php?dir=mss/628-helmst> and $\mathbf{W}_2$ at <http://diglib.hab.de/wdb.php?dir=mss/1099-helmst>.
[3] 'Imagedigitalisate von HAB-Beständen werden als Public Domain publiziert' (Wolfenbütteler Digitale Bibliothek 2013).

images 00001.jpg–00534.jpg.

Obtaining usable images for **F** was a little more complex. DIAMM links to the *Biblioteca Medicea Laurenziana Digital Repository* (Biblioteca Medicea Laurenziana n.d.), which contains no usage information, but the same images have also been uploaded by the host library itself to IMSLP as public domain (International Music Score Library Project 2013). I attempted first to extract the images from the IMSLP PDFs but, like the HAB viewer PDFs, these extracted images were of low quality and low resolution (roughly 800×1000 pixels). The *Biblioteca Medicea Laurenziana* digital repository itself serves its images through a viewer into an International Image Interoperability Framework (IIIF) server, *IIPImage*. IIIF serves images as a series of tiles which are reconstructed on the client: the entire image is never loaded at once. This is useful and bandwidth efficient when viewing large facsimile images at multiple resolutions and zoom levels, but makes the process for viewing the entire image at maximum resolution a little more involved, as the original image must be reconstructed from its constituent tiles. The browser extension "Dezoomify" can be used to find the IIIF manifest that contains the tiles for this image (Lojkine n.d.). For the first few images, Dezoomify reported a base URI of <http://mss.bmlonline.it/fcgi-bin/iipsrv.fcgi?Zoomify=\\TECA-NAS\Teca\Plutei\Lotto02\.\P000792_29.01\pyr\> and files: "Plut._29.01_0001.tif/ImageProperties.xml", "Plut._29.01_0001a.tif/ImageProperties.xml", "Plut._29.01_0001b.tif/ImageProperties.xml", "Plut._29.01_0002.tif/ImageProperties.xml", "Plut._29.01_0003.tif/ImageProperties.xml" etc. Extrapolating from this, the URI is generally a monotonically increasing four digit integer, with some additions for views of the spine and pastedowns. The final integer in the sequence was 0894. However, each IIIF manifest contains only a single line detailing how to extract the tiles of the image and, unlike the HAB viewer, no licencing information. This should not matter as, firstly, for the purposes of analysis this academic work falls under Fair use (in US law) and Fair dealing (in UK law under the specific exception for text and data mining). Regardless, there is a case to be made that facsimile reproductions of public domain works cannot be held in copyright as they are not original works.[4]

The HAB images could be downloaded on the command line by using the program

[4] In US law, this was tested in *Bridgeman Art Library v. Corel Corp* (1999) which used both UK and US law to judge that an exact photographic reproduction of Hals' *Laughing Cavalier* (1624) could not be copyrightable as the reproduction lacked originality. In UK and EU law, there is a lower standard for "originality" in images, which at times contends that "original" concerns not invention but origin. However, even where considerable effort has been made to photograph an already two-dimensional item, UK law generally holds that there is little difference between photocopying — very clearly a case of an uncopyrightable derived work — and an exact photographic copy (Stokes 2020, pp.126–141). The UK government's own guidance on digital images states that 'copyright can only subsist in subject matter that is original in the sense that it is the author's own "intellectual creation". Given this criterion, it seems unlikely that what is merely a retouched, digitised image of an older work can be considered as "original"' (Intellectual Property Office 2021).

"wget" in a loop, but these IIIF images required a more complex process, stitching the tiles together to form a large composite image. The Dezoomify browser extension provides a link to download a single image but does not provide the possibility for batch download. Downloading each image manually through Dezoomify would again be extremely laborious. However, the IIIF manifest can be used directly in a loop using the browser extension's companion program "dezoomify-rs", which parses the file and reconstructs the original image. Given the "-l" flag, dezoomify-rs selects the highest resolution tile set. The following script was therefore used to download the highest resolution images of **F** possible:

```
#!/bin/bash
# Set the base URI
url='http://mss.bmlonline.it/fcgi-bin/iipsrv.fcgi?Zoomify=\\TECA-NAS\Teca\'\
'Plutei\Lotto02\.\P000792_29.01\pyr\'
# Iterate a four digit integer from 0001 to 0894
for i in $(seq -f "%04g" 1 894); do
    # Generate the filename using that integer
    filename="Plut._29.01_${i}.tif/ImageProperties.xml"
    # Pass the result to dezoomify-rs
    dezoomify-rs -l "${url}${filename}"
done
```

7.2. Image importing

CANDR does not provide a method for the batch uploading of images into the database as facsimile instances.[5] All it provides is a user interface for adding folios, facsimiles, and uploading images to attach to facsimile records one-by-one. With thousands of images to catalogue, however, this would have been a monumental task. I therefore took advantage of the fact that the downloaded images generally followed an integer sequence, and that most of the sources' folios could be enumerated as 1r, 1v, 2r, 2v, 3r, 3v, etc. I wrote another short shell script:

```
#!/bin/bash
# Read in variables from user
read -p "Start jpeg (no zero pad): " start
read -p "Start folio: " folstart
read -p "End folio: " folend
read -p "Filename width: " filewidth
read -p "Filename extension (no dot): " fileext
# Iterate from folstart to folend
for i in $(seq $folstart $folend); do
    # Print two lines, such as:
    # 5r    00010.jpg
    # 5v    00011.jpg
```

[5]Commonly, the term "facsimile" may refer to the entire set of images relating to a source but, as shown in Fig. 4.2, a facsimile in the context of CANDR is a model of a single image, belonging to a folio and facsimile set, and containing multiple systems of music.

```
        printf \
        '%sr\t%0'"${filewidth}"'d.'"${fileext}"'\n' \
        '%sv\t%0'"${filewidth}"'d.'"${fileext}"'\n' \
        "$i" "$start" "$i" "$((start + 1))"
        # Increment the integer sequence by 2
        let "start=start+2"
    done
```

This automatic "foliator" script generates a mapping from folio name to image filename. It prompts for the starting image number, its respective folio (on recto), and the ending folio (on verso). It then prints two lines for each folio number: one with "r" for recto, and another with "v" for verso. To each it adds the next image number. There are exceptions to this scheme of course, such as inconsistent foliation, missing folios, and extra images showing details or alternate views of the manuscripts. The resulting foliation file was manually edited to reflect these inconsistencies.

A second "uploader" script uses the "curl" program to act like a scriptable web browser: it logs into CANDR using the administrator user account and password, and then for each line in the foliation file it navigates to the "add folio" page, sets the title to the folio name, clicks the "add facsimile" button, and uploads the image to the site. This way, a laborious process that would have taken weeks was accomplished in minutes.

7.3. Data input

As explained in "5.4. OMR", the input loop of CANDR is modelled after adaptive OMR: as staves are input, the OMR application is trained on this new data, and new or updated predictions are made for further staves. CANDR does not yet provide automation for the definition of system and stave quadrilaterals — although there are plans to do so in future — therefore these had to be input manually. Fig. 7.1, graphically shows the progress of the input effort that I made over June–December of 2021. The ambitious initial plan aimed to complete the input of **F**, $\mathbf{W_1}$, and $\mathbf{W_2}$ over this period. However, there were some unexpected technical teething problems in the OMR system that had to be fixed, and the demanding pace of input proved difficult to sustain. As of writing only $\mathbf{W_1}$ and $\mathbf{W_2}$ have been fully added to the system. The largest source, **F**, remains unprocessed, although initial work on some substitute clausulae was undertaken. I stopped input by a cut-off date of January 2022, as the programming toolkit for analysis was not yet developed at that time and was more intricate than initially envisioned. Despite this, the CANDR dataset currently stores 2 complete sources, 931 folios, 5,626 systems, 9,710 staves, and 560,293 notational elements. These are also organised into 896 discrete settings of music. While the dataset has achieved a level of completeness and diversity, it is worth noting that further research could potentially expand it to encompass **F** as well as peripheral sources, offering a broader scope for exploration in future.

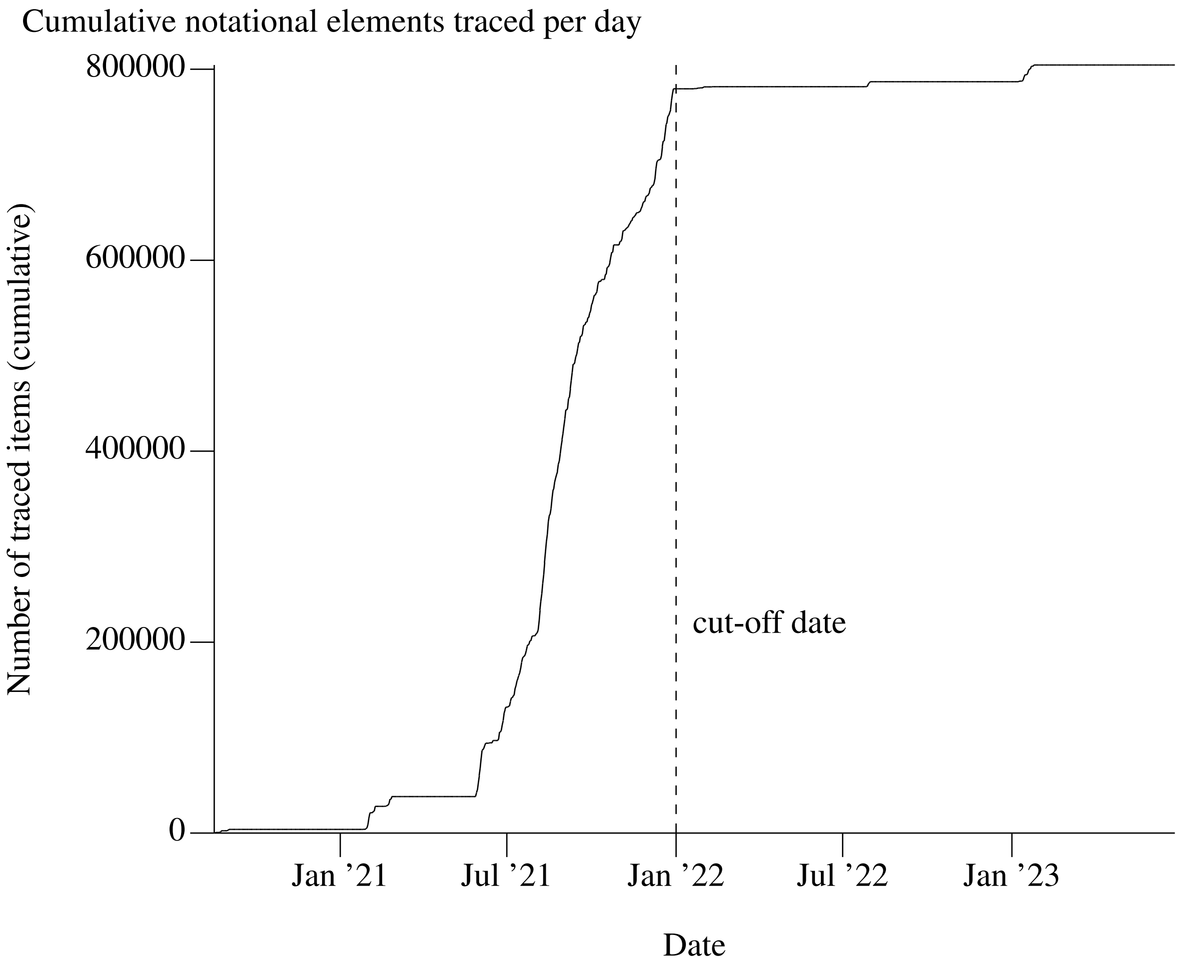

Fig. 7.1

7.4. Baseline: *Word2Vec*–DTW

To evaluate the methodologies described in "Chapter 6: Offline", a baseline analysis methodology was developed that used more conventional structures and approaches. This baseline methodology uses word embedding (*Word2Vec*) followed by the simplest alignment methodology considered (DTW), and is thus termed here the *Word2Vec*–DTW baseline methodology. This baseline will be a current state-of-the-art methodology against which the new methodology will be finally evaluated. However, there are also multiple candidates for the new methodology: three node embedding models, followed by numerous different alignment methodologies. "7.5. Node embedding: GNN" will evaluate which of the new GNN node embedding methodologies is the most effective at representing notational structure, and "7.6. Alignment: DTW or OT" will evaluate which alignment methodology yields the clearest results when applied to the node embedding models. This will result in a single "best" new methodology which will be compared to the baseline *Word2Vec*–DTW

Word2Vec-generated embedding of monophonic dataset (colour represents node type)

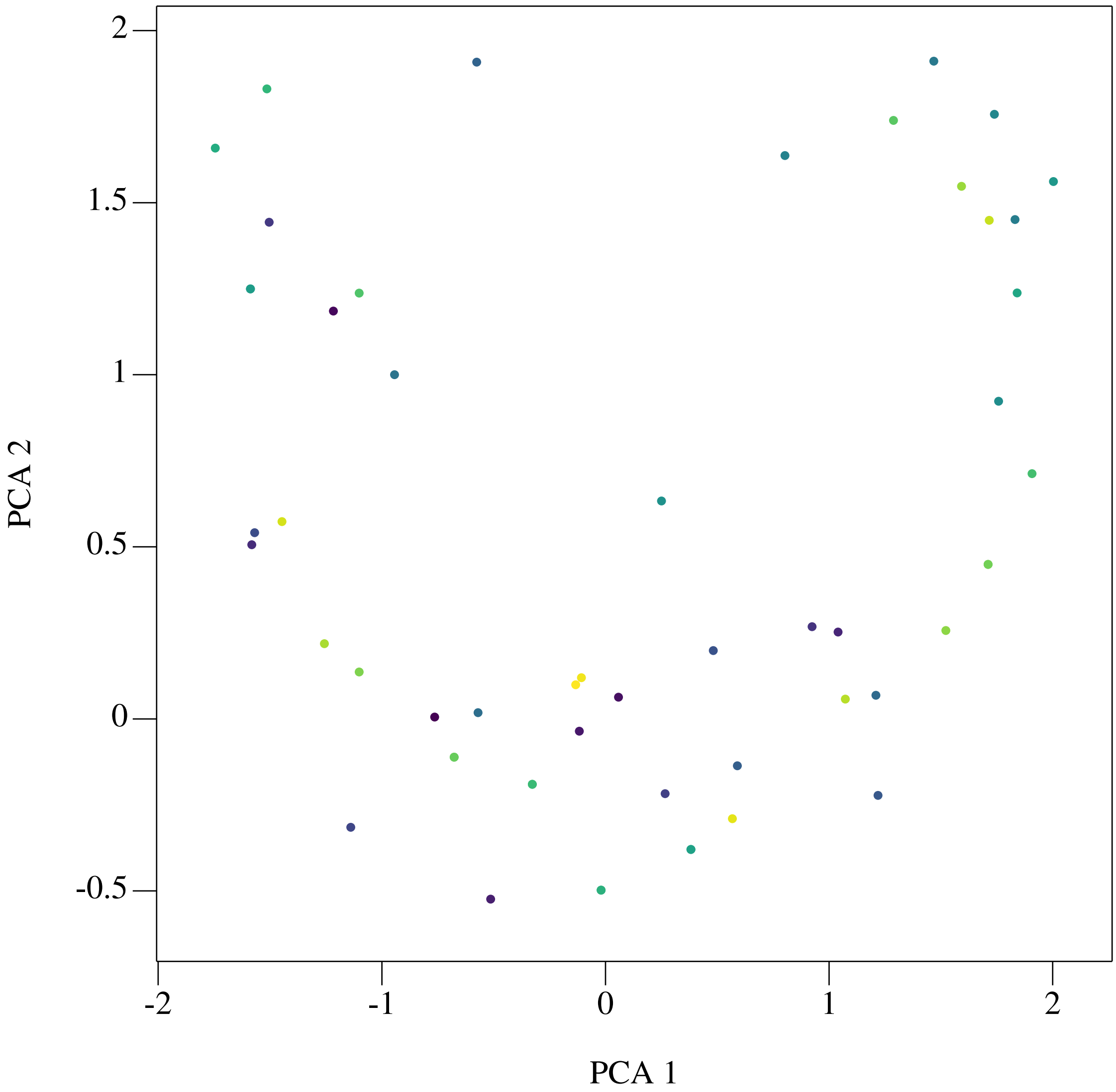

Fig. 7.2

methodology.

The baseline methodology follows a linear approach representing the current state of the art, where each line of music is extracted from its context and analysed independently. This gives two distinct advantages as a baseline. Firstly, by removing the polyphonic context in which the music exists, the baseline methodology does not have to consider the harmonic relationships between a line of music and other sounding voices. Secondly, modelling monophonic lines enables a perspective closely aligned to more traditional natural language processing (NLP) methodologies, as the musical material could be tokenised in the same way as text. However, as previously explained in "3.6.4. Encoding conclusion", this would misrepresent the music somewhat, modelling this polyphony as no different from a set of entirely independent monophonic lines, and all aspects of harmony, counterpoint, and melodic interplay would be entirely lost.

There are no current methodologies that are capable of analysing polyphony without first decomposing it into independent lines or slices as chords (see p.200). Nonetheless even with these limitations, an approach analysing each line independently can be created as a baseline with which to compare this study's chosen methodology. In terms of the overall process, the baseline methodology was therefore designed to be as similar as possible to the new methodologies, but with more standard technologies used in each component of the analysis, adopted directly from NLP. Rather than operating on notation graphs as stored in the CANDR dataset, the baseline methodology began instead by extracting all the monophonic lines from each graph and representing them as simple text files. A *Python* script converted the node attributes into unique node types, each represented by a word derived from the data structure for each node type. For example, an F at octave 3 could be represented by the pseudo-word `MeinotePnameFOctThree` which would distinguish it from an F at octave 2 (`MeinotePnameFOctTwo`). Each node was converted into a pseudo-word in this manner, and written into a series of text files where these words were separated by spaces. This enabled the text files to be tokenised and processed in the exact same way as any other text document using NLP methodologies.

Embedding was still required however, as alignments still needed to be created between independent lines. I opted for a commonly-used embedding methodology used in NLP, *Word2Vec* (Mikolov et al 2013), to generate fixed-length vectors for each word in each text file. Mikolov et al (2013) advocate for using vector sizes in the hundreds for a dataset of English but, given that there were only 48 unique words in the monophonic dataset (i.e. unique notational elements), a smaller vector size of 64 was chosen, along with a window size of 3. A *Word2Vec* model was trained over 100 epochs on the dataset. Fig. 7.2 shows the result of this embedding when the 64-size vector is projected into 2D by Principal Component Analysis (PCA). Although there are over 400,000 data points in this plot, only a few are visible as most of the same type occupy the exact same point in vector space. The points themselves are well distributed, but the space is sparsely populated with poor coverage.

The simplest alignment technique, dynamic time warping (DTW), was chosen for the baseline methodology. This is because it makes little sense to attempt optimal transport (OT) in this case as the NLP approach has already flattened the individual voices into linear sequences, and DTW matches linear sequences. OT would offer no improvement in the baseline case. DTW was therefore calculated between all pairwise sequences of tokenised monophonic lines, by using the *Word2Vec* model to convert the tokens into vectors, and then applying DTW with a cosine distance.

To evaluate the separation between similarity and difference created by the

Raw and smoothed *Word2Vec*–DTW alignment between unrelated settings

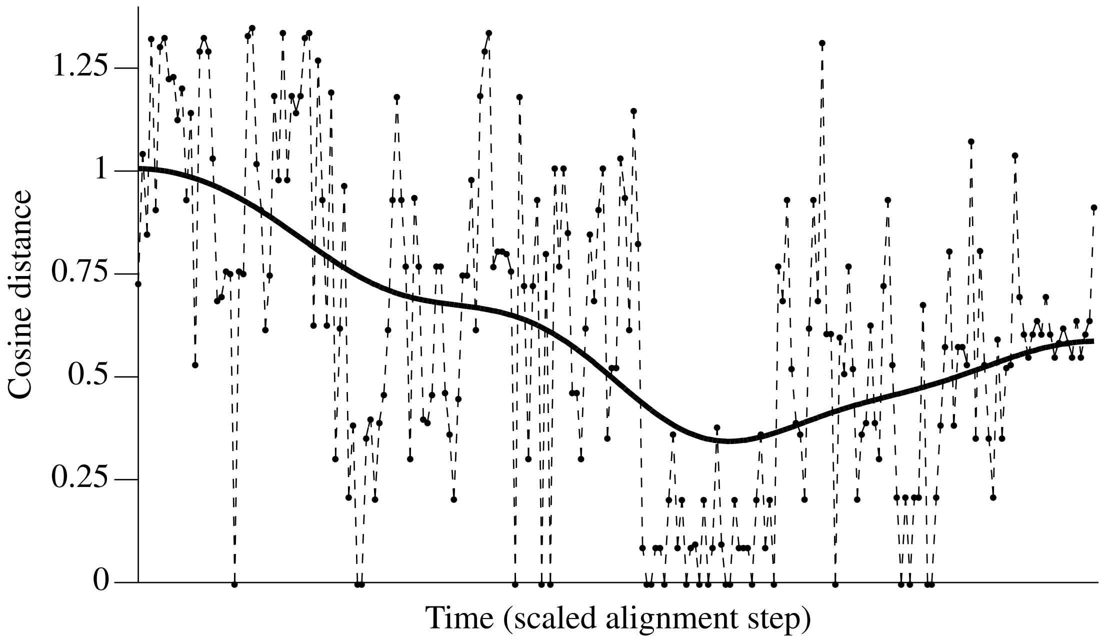

Fig. 7.3

Raw and smoothed *Word2Vec*–DTW alignment between related settings

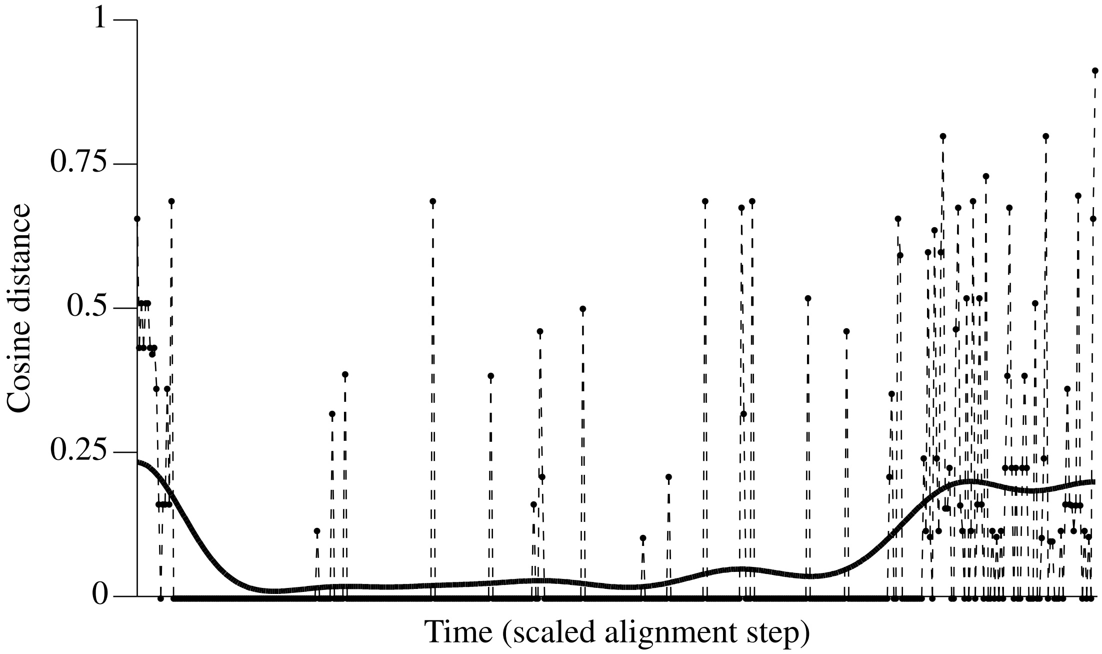

Fig. 7.4

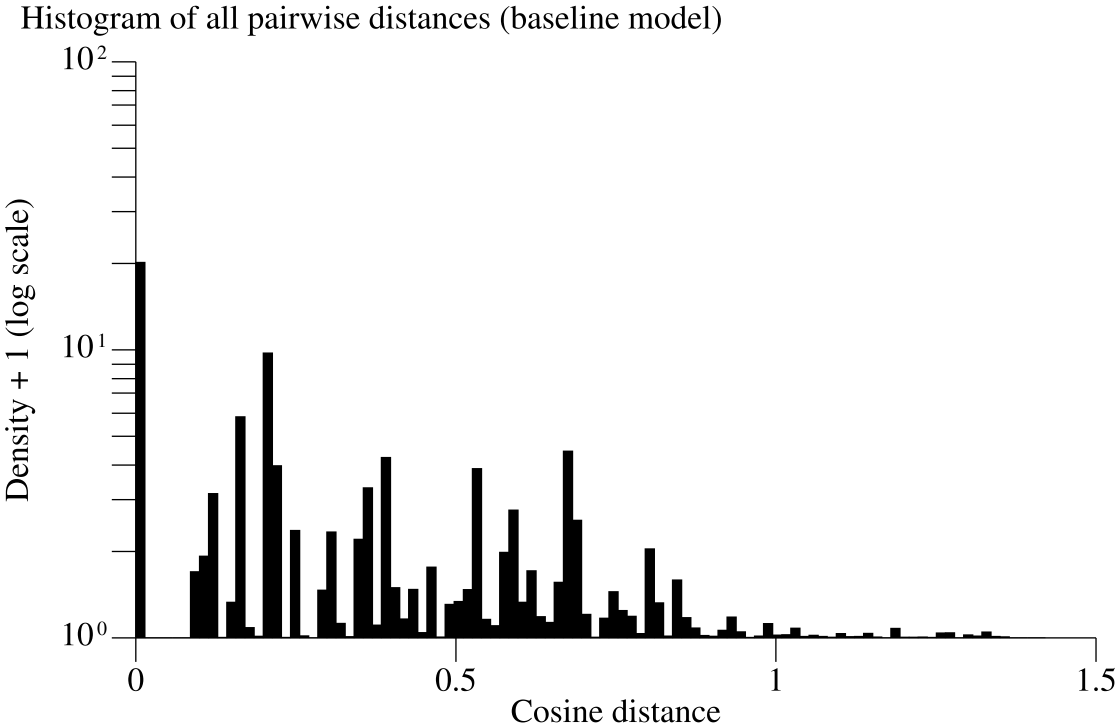
Histogram of all pairwise distances (baseline model)
10^2
10^1
10^0
Density + 1 (log scale)
0
0.5
1
1.5
Cosine distance

Fig. 7.5

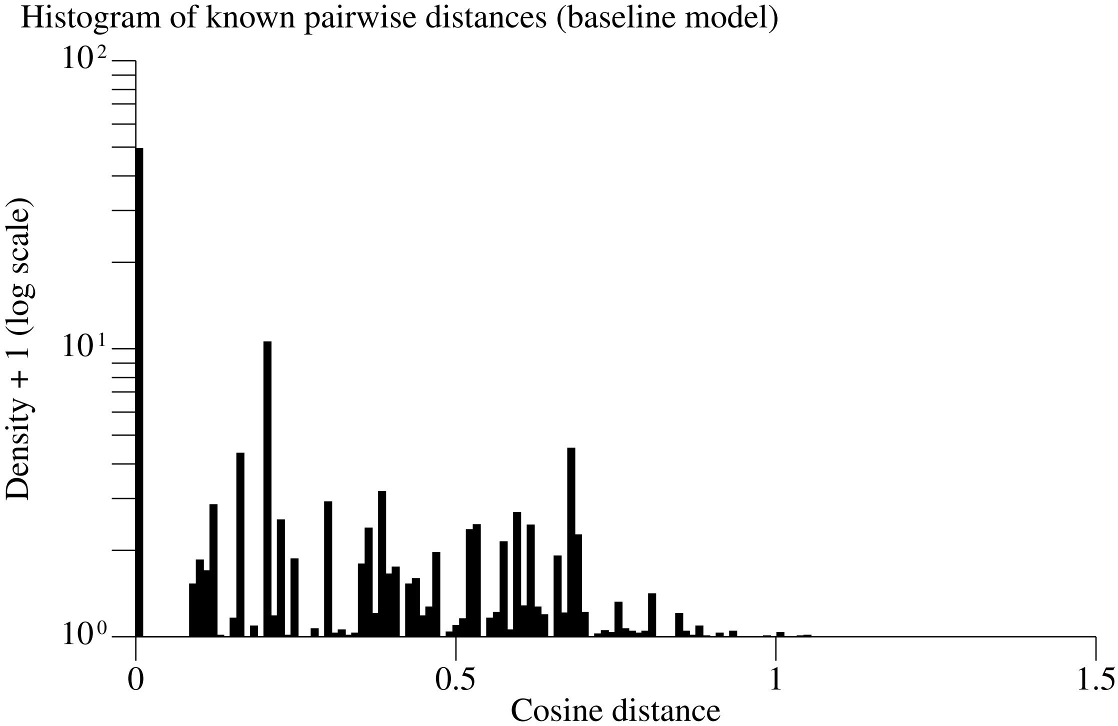
Histogram of known pairwise distances (baseline model)
10^2
10^1
10^0
Density + 1 (log scale)
0
0.5
1
1.5
Cosine distance

Fig. 7.6

Histogram of all pairwise distances (smoothed baseline model), with fitted Gamma distribution

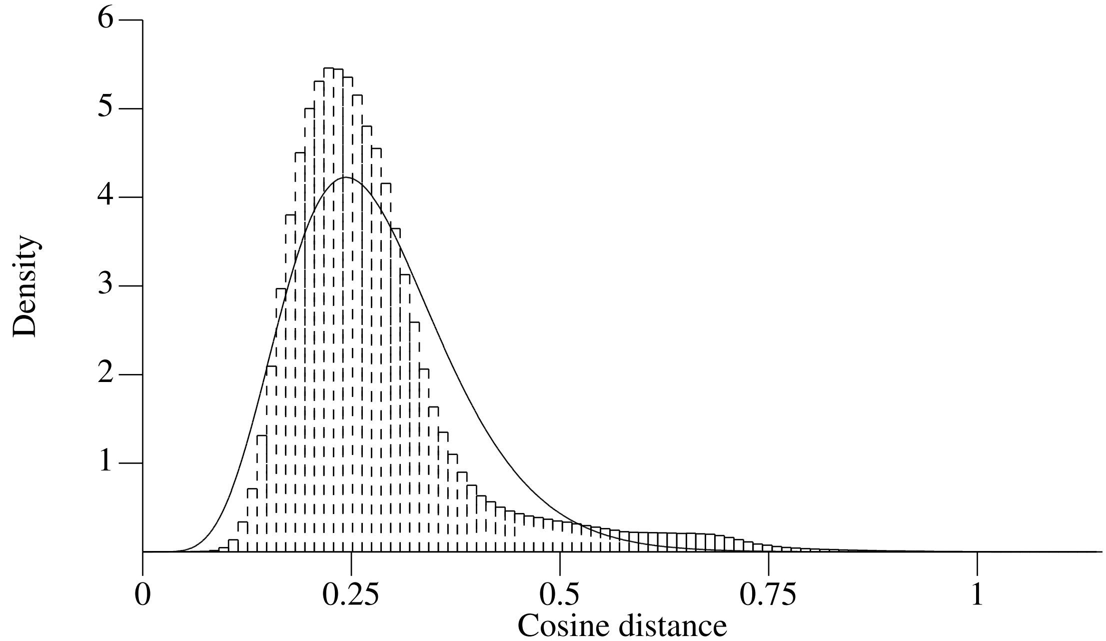

Fig. 7.7

Histogram of known pairwise similarites (smoothed baseline model), with fitted Gamma distribution

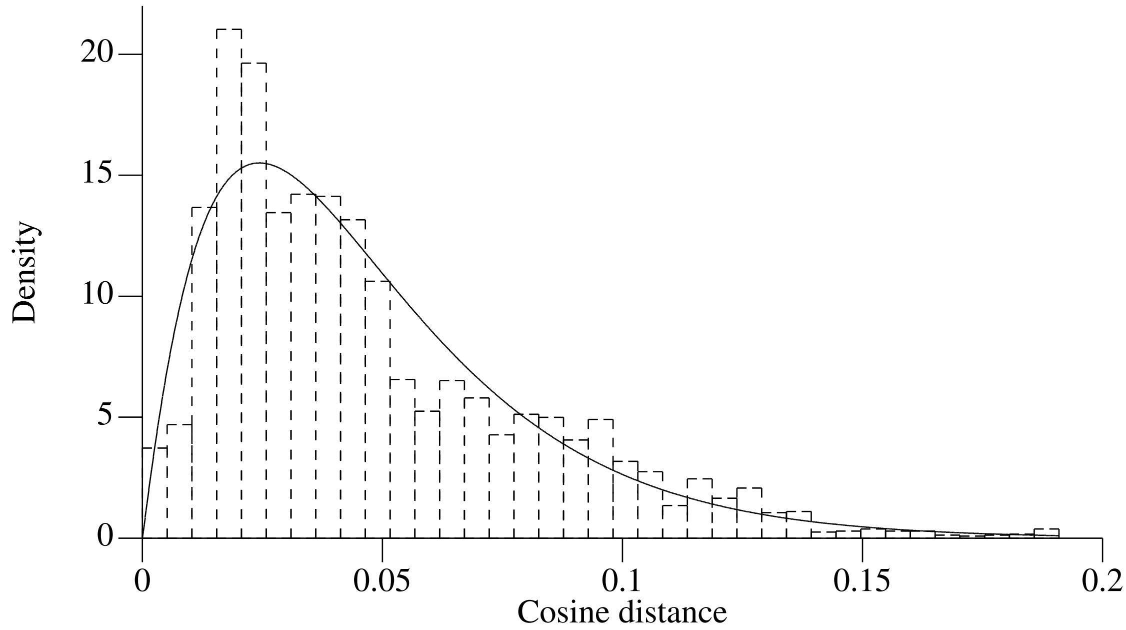

Fig. 7.8

Word2Vec–DTW methodology on the monophonic dataset, and therefore the likely success of applying a Hidden Markov Model (HMM) to these distances, a subset of the data can be extracted from known concordances within the repertory. Van der Werf (1989, pp.161–163) provides a catalogue raisonné of organa transmission, and from this can be extracted a list of organa that are transmitted in both $\mathbf{W_1}$ and $\mathbf{W_2}$. For four voices: M3. For three voices: O2, M14, M38, O27, BD-IV.2, and BD-VI. For two voices: O1, O2, O4, O5, O10, O11, O13, O24, O25, O28, O29, BD-I.1, BD-I.2, BD-I.3, M1, M2, M3, M5, M8, M9, M11, M12, M13, M14, M22, M23, M31, M32, M33, M37, M39, M42, M46, M50, M51, and M54.

The embedding and DTW process generates a signal of cosine distance against time for every pair of monophonic lines in the dataset. Figs. 7.3 and 7.4 demonstrate two results of this: the dashed line in Fig. 7.3 showing the distance between two voices likely not related (the duplum line of the conductus *Uerbum pater exhibuit*, $\mathbf{W_1}$ f.77r–v and the tenor of the offertory *Felix namque*, $\mathbf{W_1}$ f.210v–211r), and the dashed line in Fig. 7.4 the distance between two voices that are known to be related (transmissions of the duplum line of *Fuit homo* O13 from $\mathbf{W_1}$ f.20r and $\mathbf{W_2}$ f.59r). Here, the *x*-axis represents time in only a broad sense as the step from the beginning to the end of the signal. As can be seen in both cases, the signal is extremely noisy, and the distances quickly oscillate between high and low difference: either the notation matches or it does not.

This extreme matching vs non-matching pattern is reflected in the histograms of all distances (Fig. 7.5) and known similar distances (Fig. 7.6), where the large majority of distances are at or near zero. Indeed, to see any shape in the histogram we must view it with a logarithmic y-axis, and there is little distribution across the bounds of cosine distance [0, 2]. This issue can be resolved by smoothing the signal such that sudden peaks and dips are evened out. This can be seen by the solid lines in Figs. 7.3 and 7.4, which show the same signals as before, but smoothed using a Gaussian kernel with $\sigma = 20$. The similarity between the two voices in Fig. 7.4 is now much more apparent, and it is clear to see that the cosine distance stays much closer to zero than in Fig. 7.3.

This smoothing also alters the shape of the histograms in Figs. 7.7 and 7.8, which now have very few distances near zero, and peaks near 0.02 and 0.25 respectively. These smoothed histograms broadly follow Gamma distributions, and fitted Gamma probability density functions are overlaid in these figures to demonstrate this. Gamma distributions were therefore chosen as the emission distributions for the HMM, the initial state probabilities for *J* and *K* to [0.1, 0.9] respectively,[6] and the HMM transition matrix was set to:

$$\begin{bmatrix} 0.1 & 0.9 \\ 0.9 & 0.1 \end{bmatrix}$$

Samples from the “similar” subset extracted using van der Werf (1989, pp.161–163)

[6]For an explanation, see “6.8. Alignment and result extraction”.

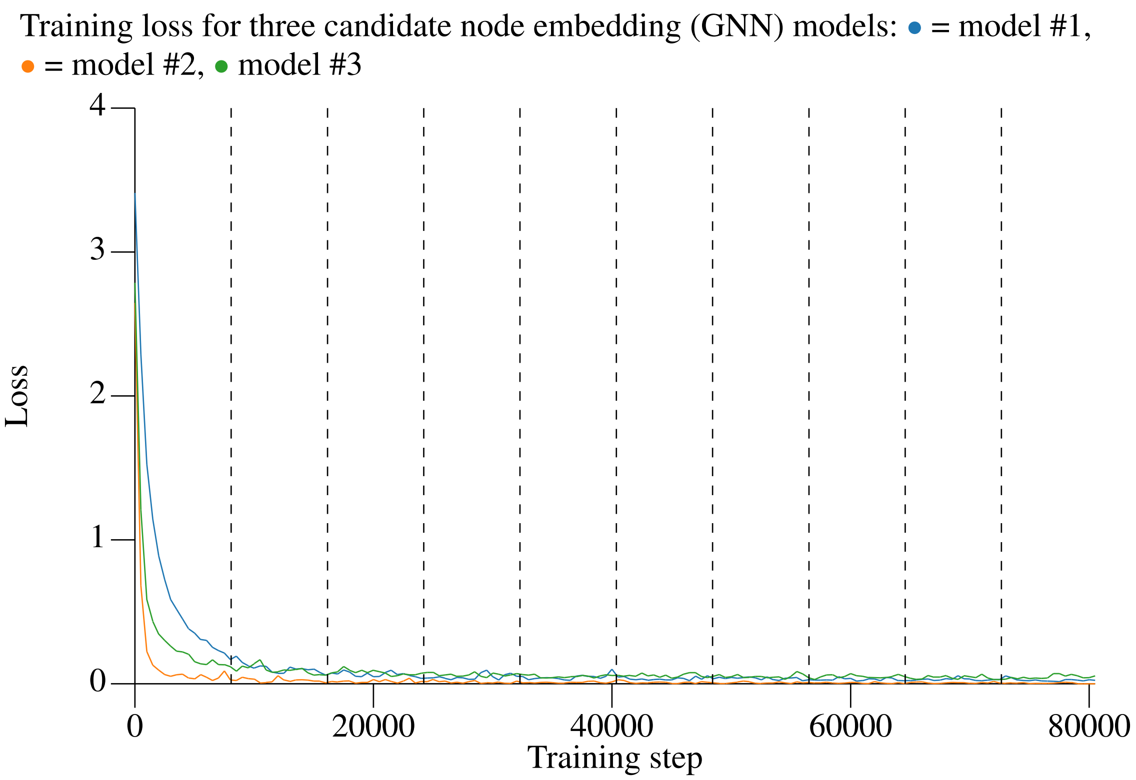

Fig. 7.9

were given a prior probability of being similar of 1.0 (i.e. definitely similar), and all other samples a prior of 0.01 (i.e. 1% likely to be similar). Gamma distributions representing *J* and *K* were fitted to the smoothed versions of the known similarity subset and the full dataset respectively. However, in training it was found that, due to the extreme imbalance between samples from both sets (approximately 1:32,000 known similarities to all distances), training the HMM directly on the full dataset caused the model to ignore the known similarity samples entirely, and always predict a "difference" state. This was resolved by randomly sampling the larger dataset into a more manageable imbalance (1:100). From these samples, each smoothed pairwise signal of cosine distance was used to train the HMM in a semi-supervised manner, using the chosen prior probabilities of each sample.

After training, the model was then used to predict the internal states ("similar" or "difference") of all smoothed pairwise samples in the dataset. Once again, the HMM was given a prior of 0.01 for each sample (i.e. 99% likely difference), and state predictions for each smoothed sample were returned. In the large majority of cases, and as expected, no similarity was predicted, returning a string of zeroes indicating a continuous predicted "difference" state. However, in areas of sustained low cosine distance, the HMM returned contiguous sets of ones within the zeroes (indicating some likelihood of a "similar" state). Some of these were false positives, especially very short sets of similarity prediction where the signal suddenly dips before returning high again, and these were filtered out by setting a minimum length of 20 for contiguous predictions. Shorter sequences were reset to zero. Pairwise distances that contained at least one set of predictions for "similarity" were plotted and saved to disk. Others were discarded. This generated 485 results which were then manually interpreted.

7.5. Node embedding: GNN

Whereas the baseline *Word2Vec*–DTW methodology began by preprocessing the notation graphs into linear voices as text files, the more advanced methodology operated directly on the notation graphs themselves, using the first four scripts described on pp.154–155 to enact the methodology as described in "6.2. Embedding": downloading MEI files from the CANDR database, parsing them into notation graphs, generating multiple ego graphs for each, and converting the data into a format suitable for the *PyTorch Geometric* library. This created a dataset suitable for training and evaluating the three different node embedding models.

Each node embedding model was trained using 80% of the data using cross entropy loss with balanced class weights.[7] For each, an embedding size of 16 was trained, using a batch size of 40, dropout of 0.2, and learning rate of 0.001 over 10 epochs. Each of these hyperparameters were selected by manually altering the parameters over a number of runs so as to minimise loss while keeping the embedding size small. The models were then evaluated using the remaining 20% in cross-validation. Fig. 7.9 shows the training loss over the ten epochs; all three models quickly learn embeddings and reach a sustained low level of loss.

For each model, an AUROC score and Calinski–Harabasz index was calculated.[8] Tbl. 7.1 presents the AUROC and Calinski–Harabasz scores for each model. A Calinski–Harabasz index is used to define the clustering of points according to their labels; a higher index indicates that points are more closely clustered. However, in this case a high index means that the node embeddings are distributed closely according to their node type, and thus the structure of the graph around the node is not having much effect. A stronger Calinski–Harabasz index is therefore a lower value. The table shows that in both AUROC and Calinski–Harabasz index the two-layer, 2nd model (the middle example in Fig. 6.2) is the best-performing node embedding model out of the three candidates, and this model will thus be selected for this study's analysis. Fig. 7.10 shows the resulting node embeddings from this model projected into 2D using PCA. Not only are they clustered by node type (the colour of each node), but in comparison to the *Word2Vec* embeddings in Fig. 7.2, they are much more evenly distributed

[7] <https://scikit-learn.org/stable/modules/generated/sklearn.utils.class_weight.compute_class_weight.html>.

[8] See p.221 for an explanation.

2D projection of GNN embeddings: each of the 43 nodes types are represented by a different colour

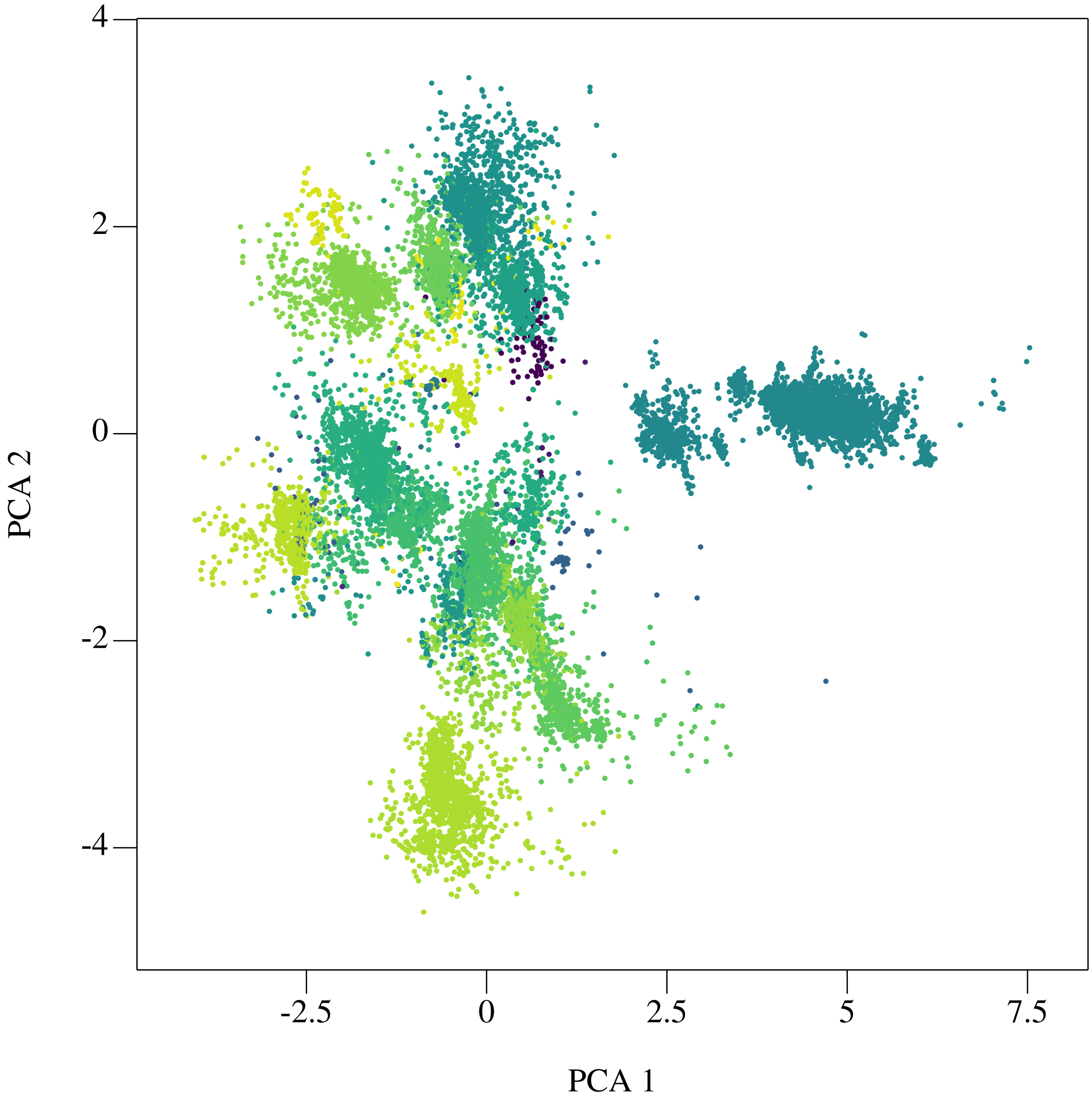

Fig. 7.10

throughout the latent space. It can be concluded from a visual inspection alone that these GNN embeddings are richer than the *Word2Vec* baseline.

Node embedding training results

Model	AUROC	Calinski–Harabasz
1	0.920	1622459
2	**0.924**	**48852**
3	0.919	67119

Tbl. 7.1

7.6. Alignment: DTW or OT

The selected node embedding model provides a way of mapping nodes in notation graphs to fixed-length vectors. Similar to the baseline methodology, pairs of notation graphs can then be aligned and distances calculated. However, unlike the baseline which opted for the simplest alignment using DTW, there are instead multiple methodologies to evaluate: as well as DTW itself there are two different positional encoding schemes for OT (temporal augmentation and RoPE). Grid search was therefore employed to find the optimal alignment methodology. As with the baseline methodology, every pair of settings is compared using the alignment technique and cosine distance, and a resulting signal is created showing cosine distance from beginning to end of the alignment. The one difference is that the new methodology makes fewer comparisons due to comparing whole settings rather than each voice independently. Once again, the same subset was extracted as training samples for similarity and the normalised histograms of both the known similarities and all distances was compared. For completeness, this was done not just for the chosen two layer model, but also the other two node embedding models (one layer and three layer) so as to further validate the selection of the two-layer model.

To ascertain which of the models and alignment methods was most successful, a variety of metrics were employed. To give the HMM the best chance of separating the emission distributions J and K, the most successful alignment methodology was deemed to be that for which the distribution of known similarities was most divergent from the distribution of all differences. These metrics were outlined in “6.8. Alignment and result extraction” as: Wasserstein distance, Kullback–Leibler divergence, and overall intersection. For each GNN model, each notation graph was mapped into vector form, and each alignment methodology was applied to the dataset in turn: DTW, temporal augmentation (TA) with $\Phi = [100, 1000, 10000]$, and RoPE with $\theta = [100, 1000, 10000]$.

The results of the grid search are shown in Tbl. 7.2. Unfortunately, selecting an alignment methodology is not as clear cut as selecting an embedding methodology, as no one metric can be said to unilaterally measure the difference between the two distributions. The furthest separated according to Wasserstein distance alone is DTW (0.25, 0.24, and 0.22 respectively), but the highest scoring for KL divergence is TA with $\Phi = 100$ and RoPE with $\theta = 100$, both for model 2. The best scoring for intersection (lowest value) is TA with $\Phi = 100$ using either model 2 or 3. Taking an overall view, it appears that many results are not separable, but the preferred OT methodologies do just as well if not better in some metrics, with higher Wasserstein distances and Kullback–Leibler divergences and lower proportions of intersection. The most performant of these appears to be model 2 using TA, with $\Phi = 100$. This model and alignment methodology

Node embedding alignment results

Model	Method	Param	Wasserstein	KL divergence	Intersection
1	DTW		0.25	0.15	0.75
	RoPE	100	0.20	0.30	0.73
		1000	0.20	0.30	0.74
		10000	0.20	0.29	0.74
	TA	100	0.19	0.29	0.72
		1000	0.10	0.29	0.69
		10000	0.02	0.23	0.72
2	DTW		0.24	0.21	0.71
	RoPE	100	0.18	0.36	0.66
		1000	0.19	0.35	0.66
		10000	0.19	0.35	0.67
	TA	100	0.17	0.36	0.65
		1000	0.08	0.32	0.67
		10000	0.01	0.22	0.72
3	DTW		0.22	0.20	0.70
	RoPE	100	0.16	0.34	0.67
		1000	0.16	0.34	0.68
		10000	0.16	0.34	0.68
	TA	100	0.12	0.35	0.65
		1000	0.03	0.27	0.69
		10000	0.00	0.03	0.94

Tbl. 7.2

was therefore chosen for the remainder of the analysis, and is termed the new GNN–OT methodology.

Using this alignment technique, we can view the differences between settings as a signal. Similar to how Figs. 7.3 and 7.4 demonstrated the resulting alignment distances for the baseline model for both a pair known to be different and a pair known to be similar, Figs. 7.11 and 7.12 present the same pairs of settings, but this time using the GNN–OT methodology. Once again, the x-axis represents time in only a conceptual sense of the signal progressing from beginning to end, not any real sense of musical time. There is an even starker difference here compared to the baseline model: yet again both signals are extremely noisy (dashed lines), but smoothing with $\sigma = 20$ (solid lines) evens out the signals and shows a clear difference between the characters of the signals. This can be seen in the histograms of the raw distances (Figs. 7.13 and 7.14) which do not follow a clear distribution, but adding smoothing causes both to closely follow a Gamma distribution (Figs. 7.15 and 7.16).

Once again therefore, an HMM was constructed with two Gamma emission distributions. As before, the initial state probabilities for J and K were set to [0.1, 0.9], and the HMM transition matrix:

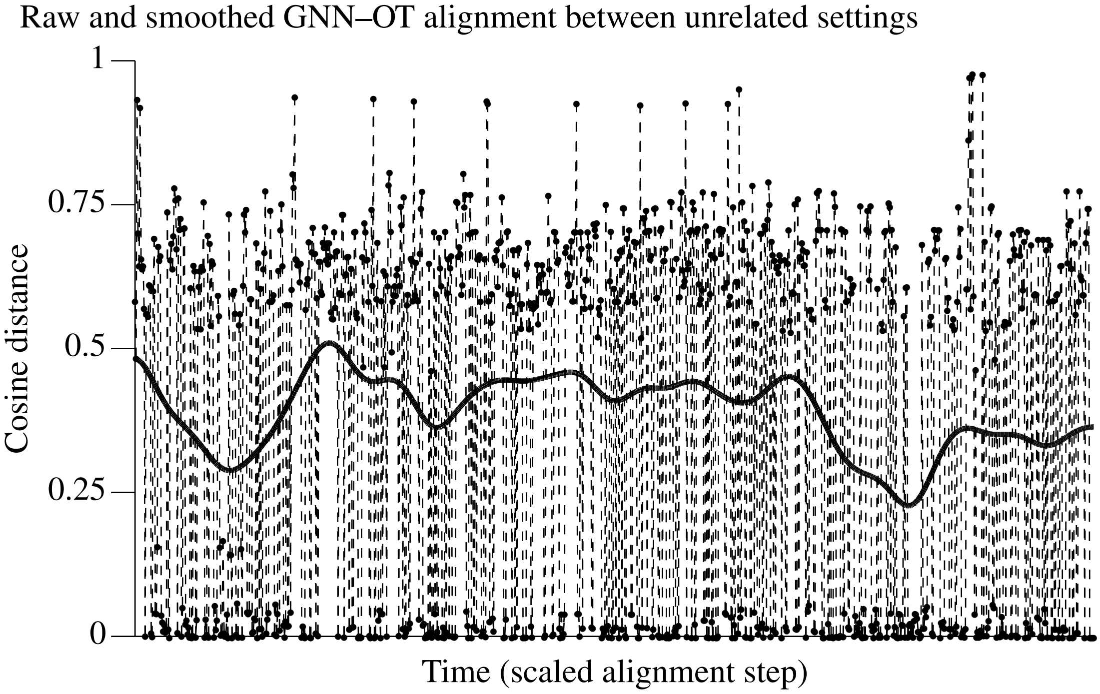

Fig. 7.11

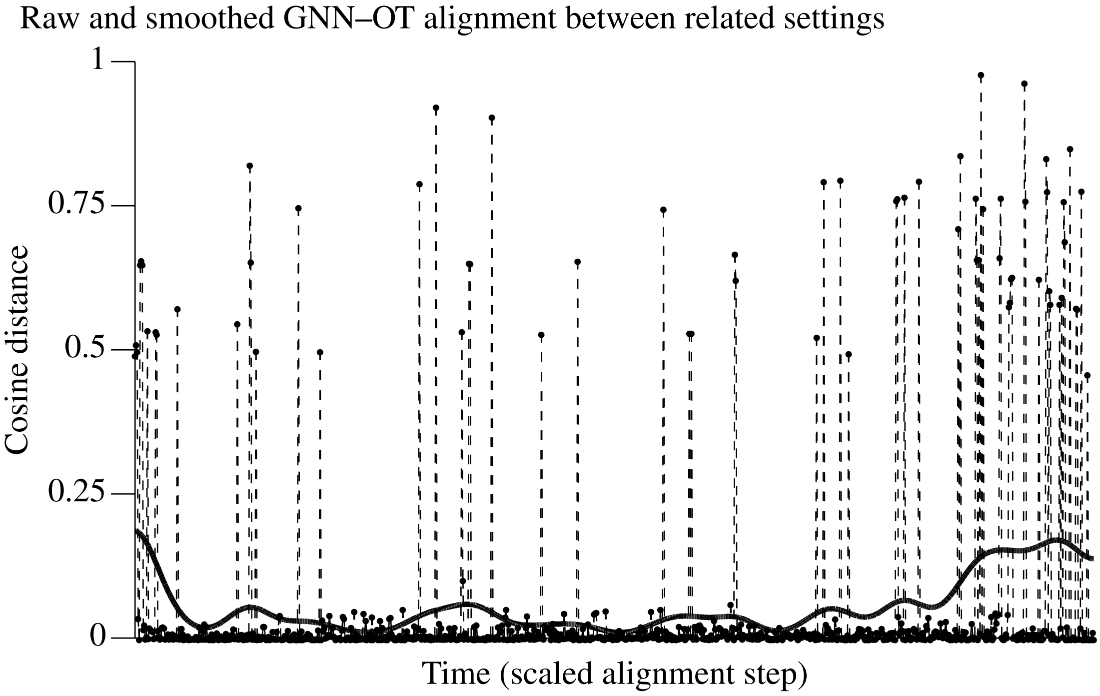

Fig. 7.12

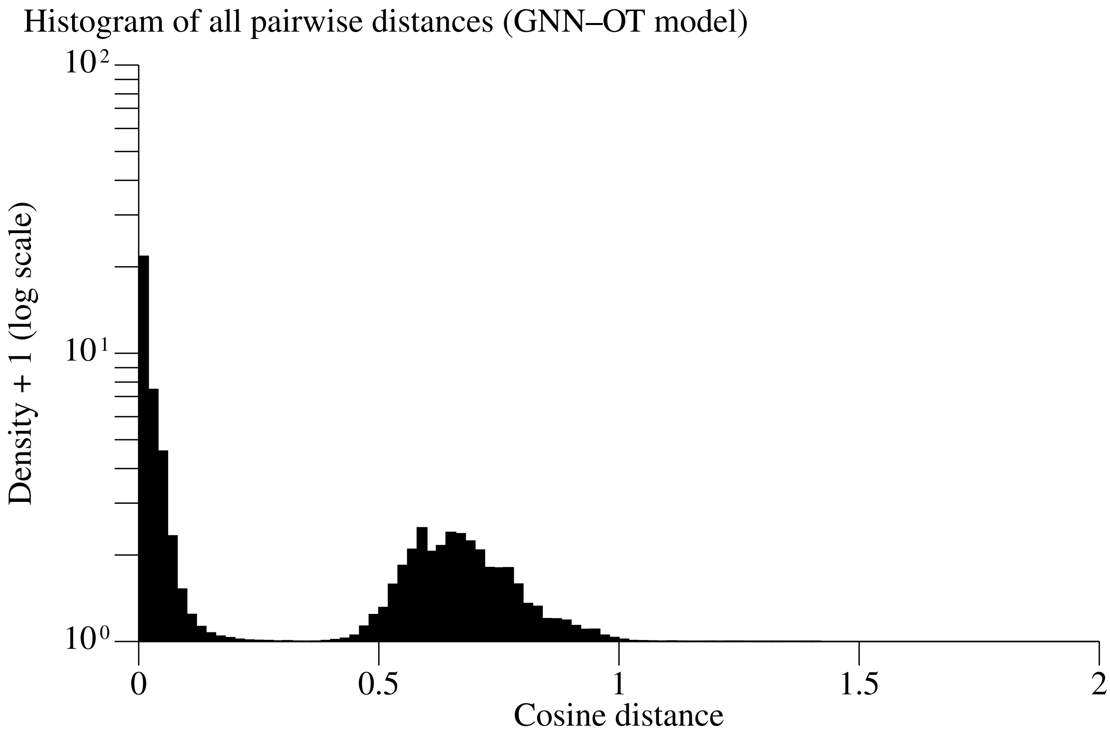

Fig. 7.13

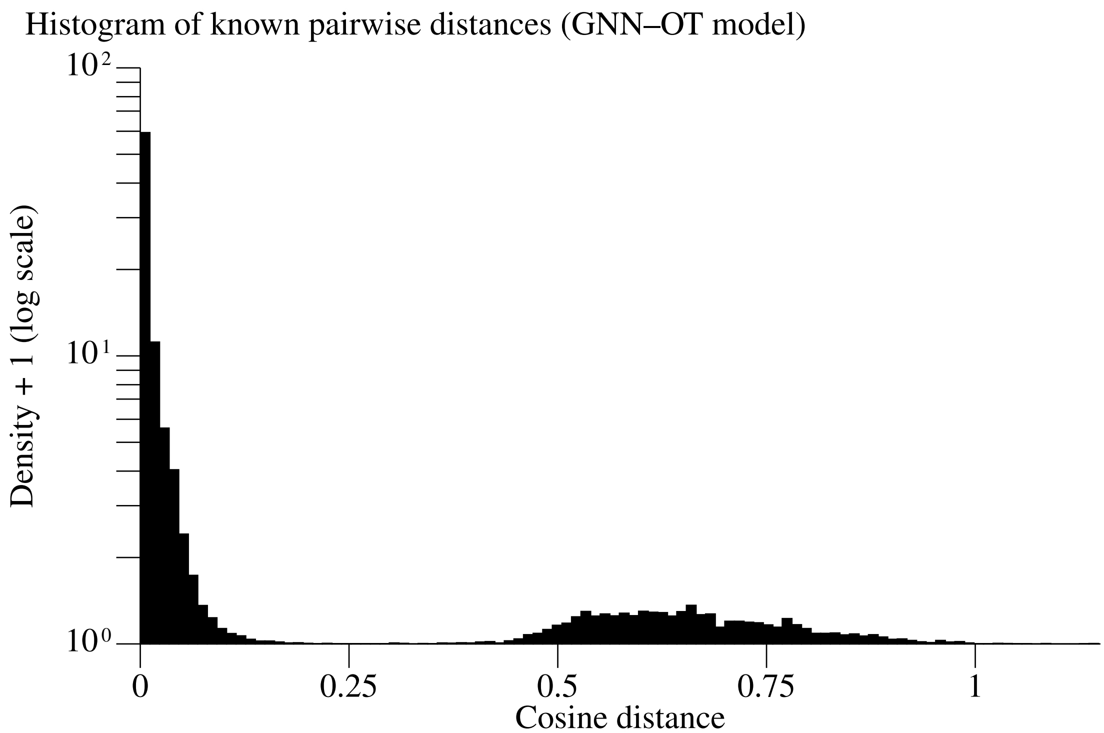

Fig. 7.14

Histogram of all pairwise distances (smoothed GNN–OT model), with fitted Gamma distribution

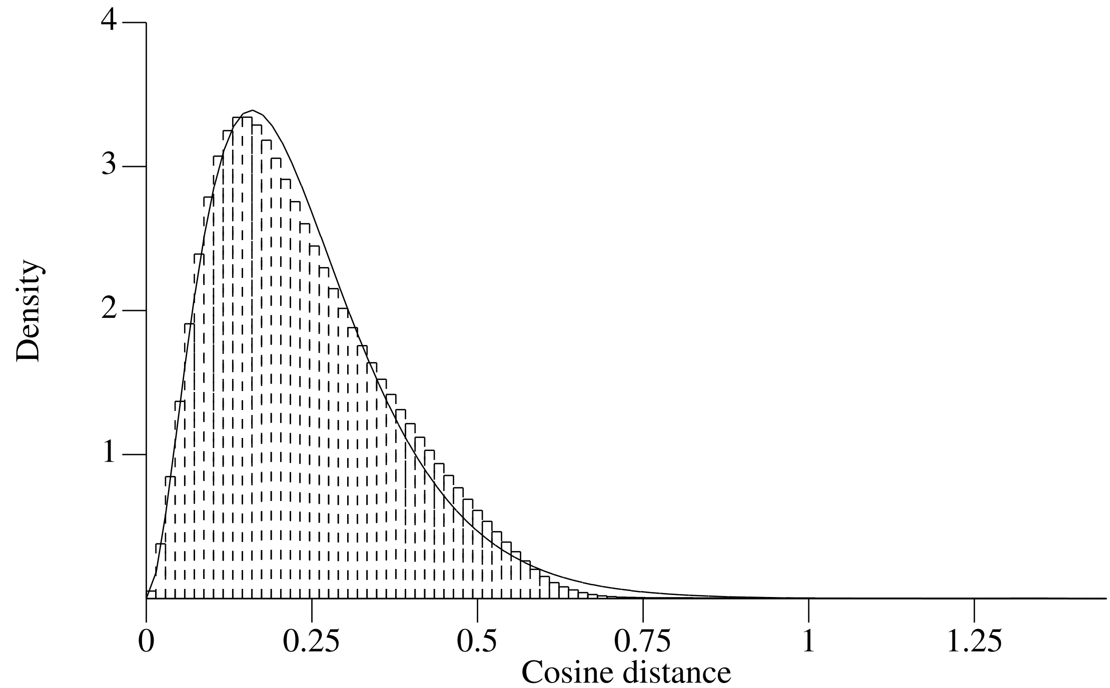

Fig. 7.15

Histogram of known pairwise similarites (smoothed GNN–OT model), with fitted Gamma distribution

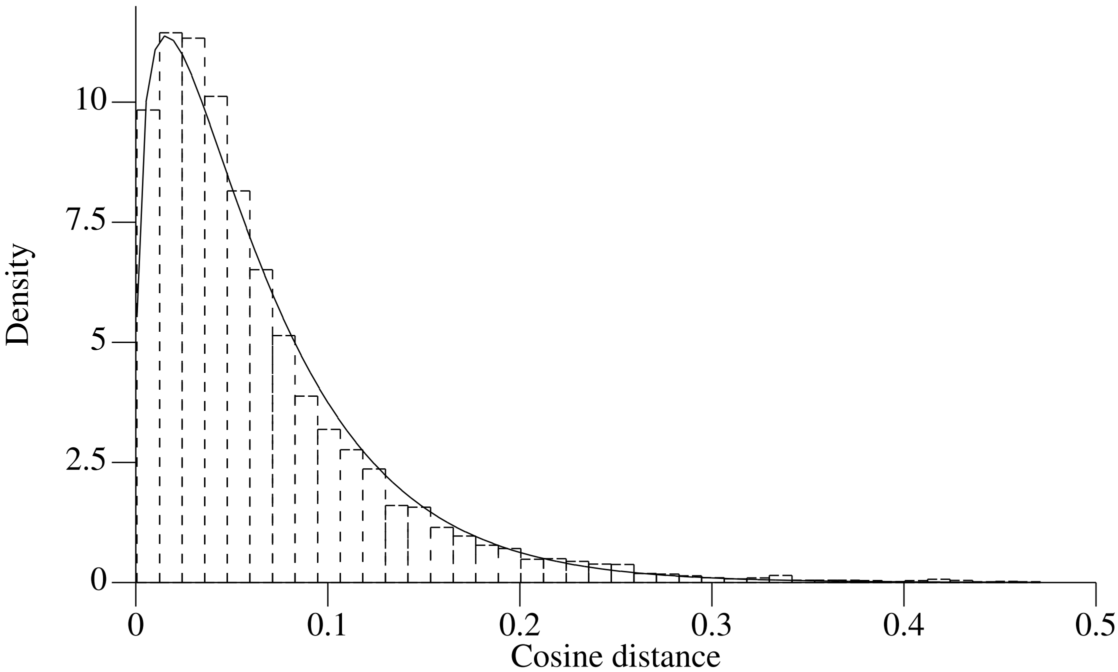

Fig. 7.16

Visualisation of results proportion for both methodologies

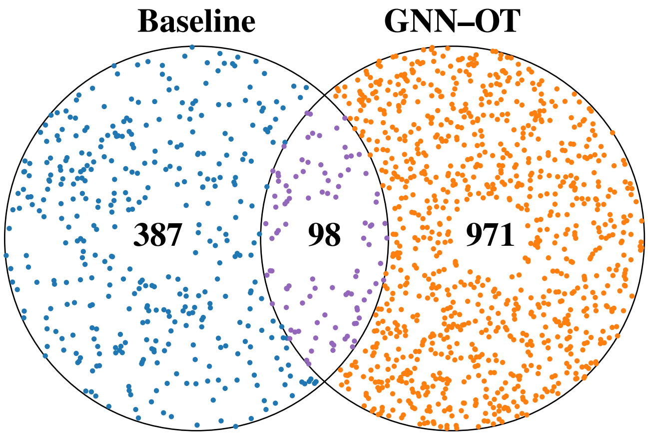

Fig. 7.17

$$\begin{bmatrix} 0.1 & 0.9 \\ 0.9 & 0.1 \end{bmatrix}$$

Samples from the "similar" subset were given a prior probability of 1.0, and all other samples a prior of 0.01. Gamma distributions were fitted to the smoothed versions of the signals, and the samples rebalanced to 1:100. The HMM was trained on these smoothed signals, predictions were extracted for all pairwise distances, and filtered using the same contiguous function as used for the baseline model. 1,069 results were saved to disk, which were then manually interpreted.

7.7. Results

Both *Word2Vec*–DTW and GNN–OT methodologies created a large number of results, and the only way in which these can be interpreted further is to manually examine each set. Fig. 7.17 represents the number of these results visually in the form of a Venn diagram: not only does the baseline methodology generate 485 results and the node embedding methodology 1,069, but 98 of these results were found to be in common between both approaches, demonstrating that a small, but not insignificant, proportion of the results were found regardless of the methodology used. However as will be seen in "7.7.1. Baseline results", the results from the baseline methodology fail to tell us much that is new about the repertory.

As an example of this intersection, we can examine one of the results found in common between methodologies: a transmission of *Ego rogabo* O10 ($\mathbf{W}_2$ f.49v–50r) and *Repleti sunt* O11 ($\mathbf{W}_2$ f.50v–51r). Although both O10 and O11 were used to train the HMM in the similarity subset based on van der Werf (1989, pp.161–163), the similarities drawn there were transmissions of the same setting, that is, similarity

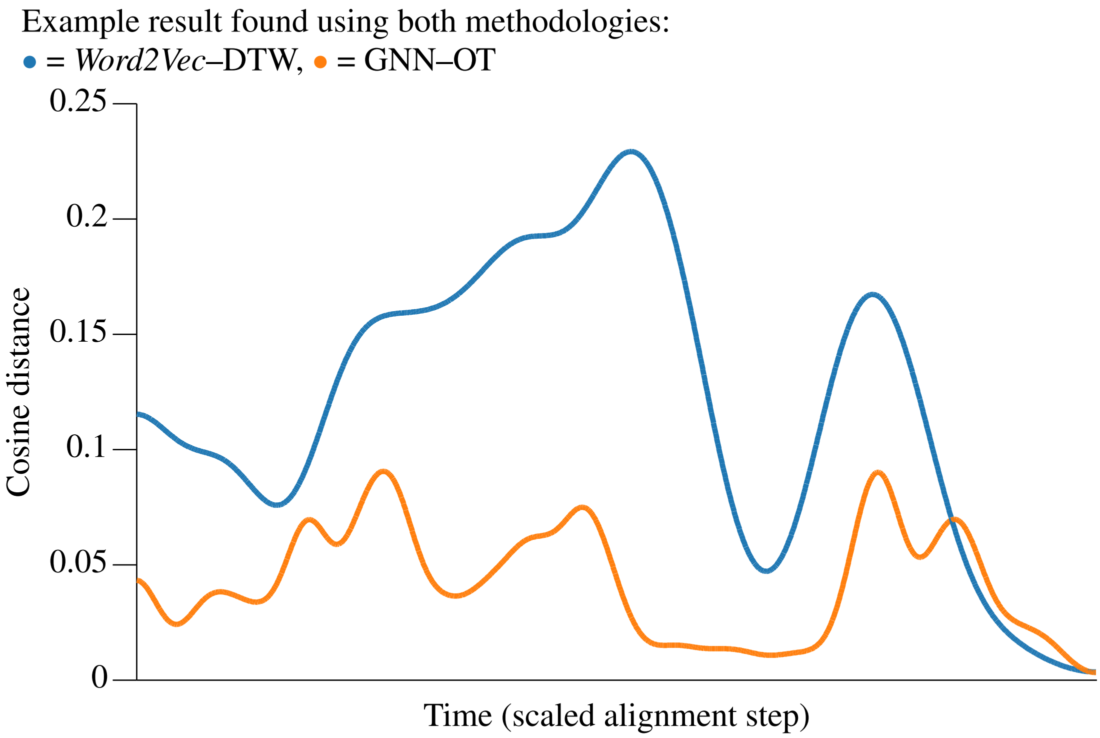

Fig. 7.18

between two transmissions of O10 and two transmissions of O11, considered independently of each other. The result shown in Fig. 7.18, however, is the similarity between the transmissions of O10 and O11 in $\mathbf{W}_2$. Their similarity is of course well known: for example, Tischler (1988, p.44) begins his discussion of these settings by stating that they are 'closely related'. What is important here is not that they are related per se, but that the result was found without the models being pretrained on this fact; the computer has detected this similarity "on its own". If the model can detect similarities such as this without being trained on this data, then it is likely that more such similarities may be found. Moreover, Fig. 7.18 shows clearly that the GNN–OT methodology is much more successful in demonstrating this known similarity. The median values for "similarity" and "difference" for both the *Word2Vec*–DTW and GNN–OT models are roughly 0.02 and 0.25 respectively (see Figs. 7.7, 7.8, 7.15, and 7.16), and although the *Word2Vec*–DTW model varies between zero and 0.23, indicating some level of similarity, the GNN–OT methodology yields a result staying quite close to zero, never going above 0.1, and therefore indicating a high level of similarity throughout.

7.7.1. Baseline results

I manually inspected all 485 results generated by the baseline *Word2Vec*–DTW methodology. It was found that, in all but a few cases, the results replicated already-known concordances within the repertory. There are far too many to enumerate them all

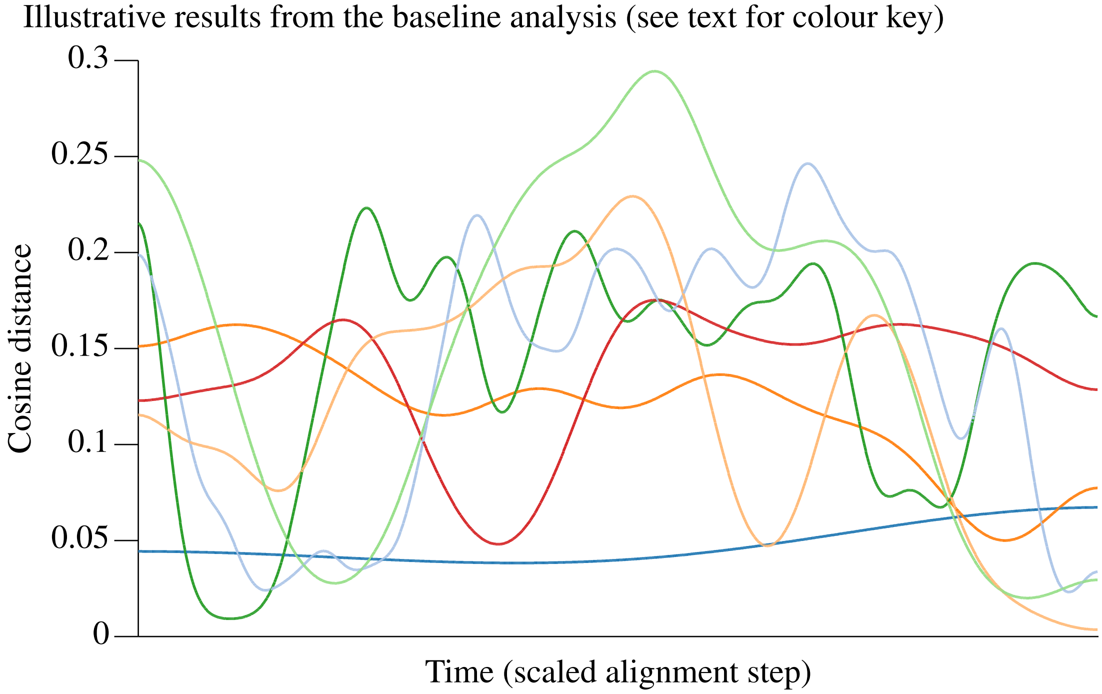

Fig. 7.19

here, but a few illustrative examples will suffice.

An important point to state first is that each result was gained by comparing individual voices only, in all possible combinations: i.e. every voice of every setting was compared with every other voice of every other setting. Therefore, a significant proportion of the results are concordances that have little connection to the polyphonic context in which they sit. 147 of the results were found simply to be concordances between tenors. For example, Fig. 7.19 (● line) shows the similarity between two motets: *Cist brunez ne me moine mie—In seculum* (**W**$_2$ f.224v–225r) and *Douce dame en cui dangier—In seculum* (**W**$_2$ f.229r–229v). The distance between the settings is not high by any means, but takes a sudden dip near the end. Rather than this being the prediction of any "real" similarity between the motets, the sudden dip can be entirely attributed to the transmission of the tenor (both *In seculum*) at the end of the setting. Although the model has detected musical reuse — use of the same tenor cannot be denied as such — it is a form of reuse that does not rely upon the polyphonic context in which it sits. Another example of this can be found in Fig. 7.19 (● line), which stays very low throughout (approx. 0.05), detecting a similarity between two *Lux magna* substitute clausulae tenors simply due to the tenors' identical rhythmic treatment; the discant of the clausulae are entirely different.

More illustrative than this are some of the other examples in Fig. 7.19. The ● line plots the difference between the two transmissions of *Assumpta est Maria* (M33) in **W**$_1$ (f.39r–39v) and **W**$_2$ (f.78r–78v). The curve of the plot follows Smith's (1964)

Smith's (1964) diagrams of M22 and M33

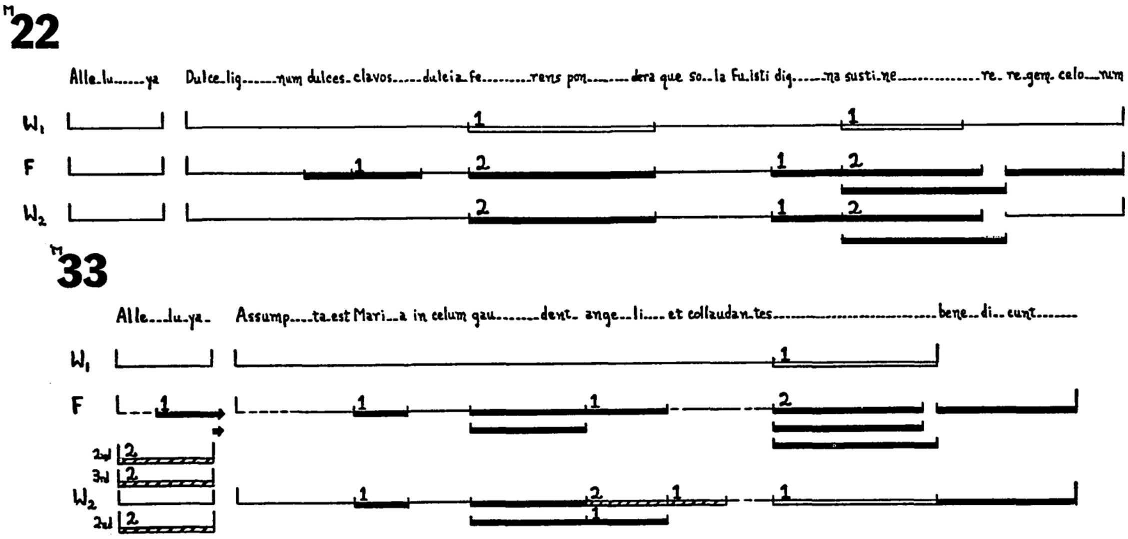

Fig. 7.20

sectionalisation of this setting (see Fig. 7.20): $\mathbf{W}_1$ and $\mathbf{W}_2$ begin *Assumpta est* with almost identical purum (corresponding to the initial dip), but then for *Maria*, $\mathbf{W}_2$ has discant whereas $\mathbf{W}_1$ continues in purum (plot rises significantly). There is a dip again at *in celum*, where both return to purum, but then the distance rises again for an extended period of divergence (*gaudent angeli et collau-*). They align again for the clausula on "tes", but finally it rises again where $\mathbf{W}_1$ cuts off earlier than $\mathbf{W}_2$ at *benedicunt*.

The ● line shows a smaller and indirect concordance between the *Gloria patri* of *Concede. Adiuuent nos* (O24), $\mathbf{W}_2$ f.55r, and the *Gloria patri* of *Ex eius tumba. Cateruatim ruunt* (O25), $\mathbf{W}_2$ f.55v–56r. These are settings of the same tenor melody, but the organum is different. However, at the text *patri et*, these two *Gloria*s temporarily align and are similar for a few ordines. This is mentioned by Payne (1996, p.351) deep in his critical apparatus to $\mathbf{W}_2$, but this plot shows this result graphically by clearly showing a dip in the middle of the signal.

Furthermore, these plots are capable of not just showing a binary distinction between "similarity" and "difference", but also creating a measure of how similar or different two settings are. The ● line demonstrates this by improving upon Smith's (1964) diagram of *Dulce lignum* (M22). As can be seen in Fig. 7.20, Smith's diagram shows $\mathbf{W}_1$ (f.34r–34v) and $\mathbf{W}_2$ (f.73r–73v) as similar for *Dulce lignum dulces clauos dulcia*, then diverging into two different discant sections for *ferens pondera*, reconverging for *que sola fuisti*, different for *digna sustinere*, and finally the same for *regem celorum*. This is useful information, but categorises similarity into either "similar" or "different". On the other hand, although the ● line broadly reflects this pattern — generally low to start, then high, dipping slightly, returning high, and ending low — the level of high and low is never the same, indicating varying amounts or levels of similarity. The first and last similarities are very close, only around 0.05 distance, and the final difference is very high, peaking at around 0.25. However, the central similarity — *que sola fuisti* — does not score as low (and therefore as similar) as the other similarities. This is because, although we can draw some similarity between the two transmissions of *que sola fuisti*, they are not identical by any means and contain a number of small differences such as added, removed, and re-ligated ordines. These subtle changes can clearly be seen by Tischler's (1988, pp.1009–1010) extremely busy edition of this passage, however the plot allows us not only observe this divergence graphically but also quantify the level of similarity in a single, easily-interpretable number.

Finally, the ● and ● lines demonstrate the scale of this process and how it is entirely agnostic — and therefore perhaps free of bias — towards any pre-existing relationships between voices. The ● line shows a similar looking plot to the ● line, alternating between similar and different sections, but this is in fact the same plot of difference between O10 and O11 already shown in Fig. 7.18. As mentioned, these two settings are very similar as they are based on the same tenor melody; they share ordines and even whole passages. However, the two pieces are difficult to collate together precisely due to their complex interrelationship. This computational methodology, on the other hand, managed to align the organal parts and calculate areas of similarity and difference in a fraction of a second for each pair of settings. We can instantly see where to look in each setting for similarity. Simultaneously, the ● line plots the difference between this same $\mathbf{W}_2$ transmission of *Repleti sunt* and the related *Gloria patri*: again they are based on much of the same material which might take hours to properly collate together by hand, but the computational methodology allows us to very quickly build whole networks of interrelationships of musical reuse simply by matching areas of high similarity together. It would not matter if these similarities occurred next to each other in the same source, or at opposite ends of entirely different sources, they would still be matched as exhaustively.

7.7.2. GNN–OT results

Similarly, I inspected all 1,069 results created by the node embedding methodology. It was found that these results reflect a combination of: a) already-known surface concordances within the repertory the likes of which the baseline methodology was capable of detecting; b) general aspects of notational and musical style; and c) possible signs of deeper similarities between settings than those that can be seen by the naked eye. Once again, there are far too many results to present each individually, but a few are selected here as indicative of the general trend within this result set. Due to each signal aligning not only one voice but all voices within a setting, the signals are in

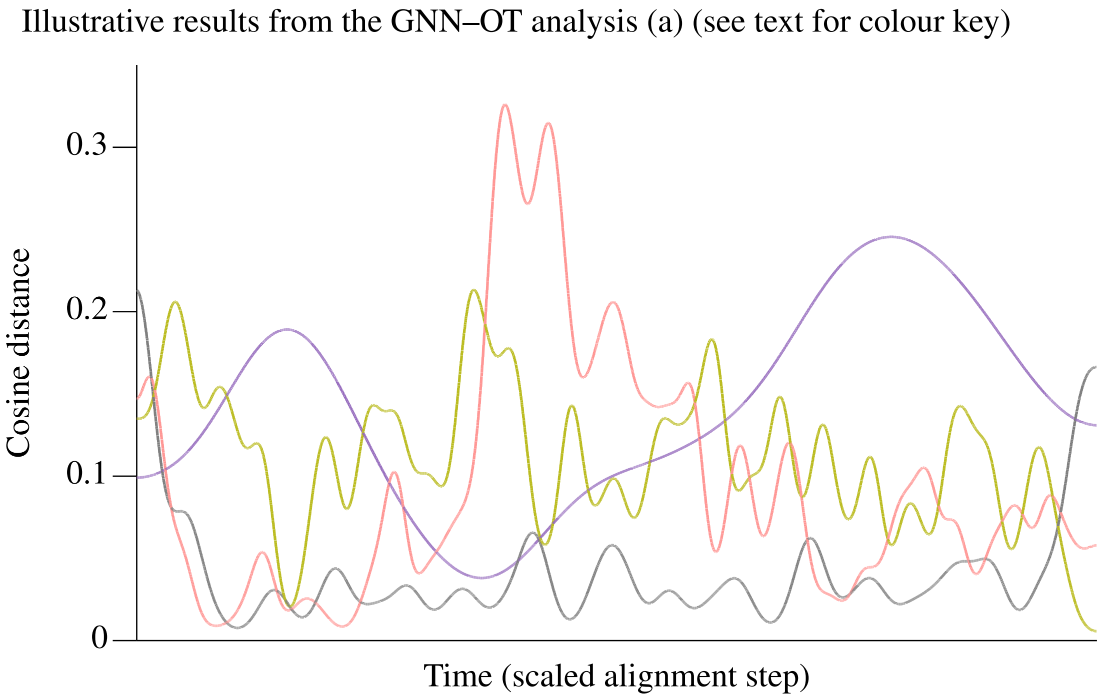

Fig. 7.21

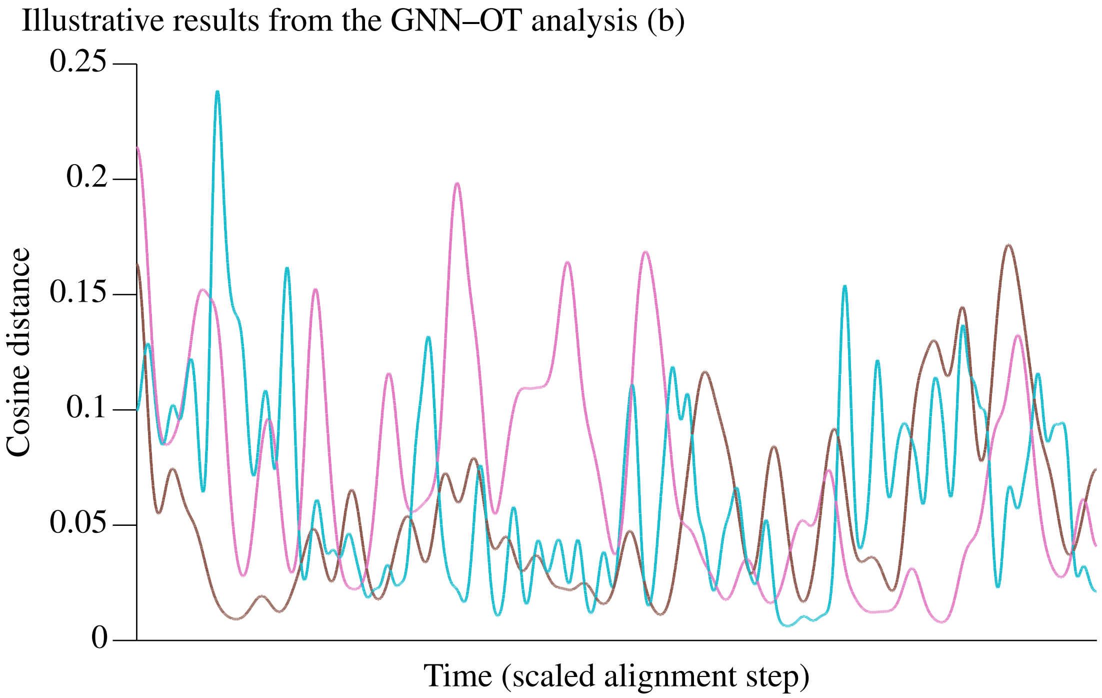

Fig. 7.22

general much longer and therefore busier than those extracted using the baseline methodology. For clarity between overlapping signals, they have been partitioned into two separate figures to minimise crossing between signals (see Figs. 7.21 and 7.22).

As has already been seen in Fig. 7.17, roughly 20% of the results generated by the baseline methodology were also found by the GNN–OT methodology, and these results can therefore immediately be explained by the same reasoning that was applied to the baseline results. This subset of results in particular demonstrates those obvious similarities already catalogued in indirect concordances and lists of clausulae. However, an immediate advantage of the GNN–OT methodology over the baseline is that it demonstrates these similarities much more efficiently and in a more holistic manner: whereas the baseline methodology generated results for only one voice at a time without considering the alignment between voices, the GNN–OT methodology is capable of analysing entire settings, including the polyphonic context in which the voices sit, and incorporating not only the linear sense of melodic contour but also the harmonic alignment between voices.

The evidence base for this overlapping set of 98 results is therefore much stronger in the GNN–OT methodology and, as was seen in Fig. 7.18, at a much more convincing signal level. For example, both methodologies detect a similarity between *Alleluya. Ascendens christus* (M23), $\mathbf{W_2}$ f.73v–74v, and *Alleluya. Iudicabunt sancti* (M42), $\mathbf{W_2}$ f.86v–87v. This similarity is already known, having been catalogued by Tischler (1988, pp.63, 71–72). In Figs. 7.21 and 7.22, the • line plots this similarity using the baseline methodology, and the • line using the GNN–OT methodology. Both plots show a low cosine distance between the two settings — and therefore evidence for similarity — but whereas the • line detects this similarity between tenor voices only, the • line demonstrates the similarity in both organal and tenor voice at once. The GNN–OT methodology therefore detects the similarity between both settings according to both voices, despite the numerous small differences between the organal voices.

Beyond this, the GNN–OT methodology also detects more distant relationships than the baseline, such as cross-source relationships and similarities beyond exact concordances. For example, not only does it detect the obvious similarity between the two transmissions of both the three-part *Sancte Germane* ($\mathbf{W_1}$ f.9r, $\mathbf{W_2}$ f.10r–10v) and *O Sancte Germane* ($\mathbf{W_1}$ f.9v–10r, $\mathbf{W_2}$ f.10v–12r), but it also detects the similarity between the respond in $\mathbf{W_2}$ and the verse in $\mathbf{W_1}$ (line •), perhaps by virtue of their similar use of melodic material. Both settings are notable within the repertory for their use of repeated-note figures and constant return to a D-A-D sonority. However, rather than simply relying on the 'intuitive statistics' of our own knowledge to make these claims (Neuwirth and Rohrmeier 2016), we can use the plot of the • line here to make the argument that this use of similar material is in fact significant within the corpus.

However, it is difficult to re-express these computational results directly in musicological terms, as the result does not reflect any exact sequence within the notation. Whereas the baseline methodology used DTW — which always keeps items strictly in sequence during its alignment process — the GNN–OT methodology uses the temporal augmentation strategy described in "6.3. Alignment" alongside OT. This augmentation methodology has no inherent constraints to keep items in order: it is only

Out of sequence match between *Sancte Germane* and *O Sancte Germane*, according to the GNN–OT methodology

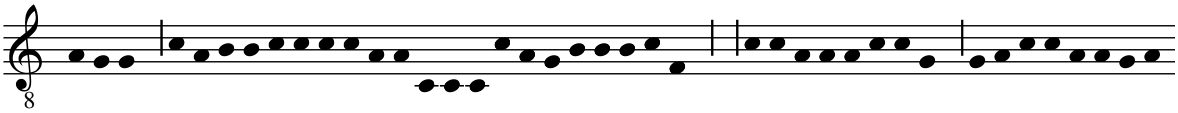

Fig. 7.23

the scaling parameter (Φ) chosen that will cause items aligned from further away to be marked with a higher penalty. This methodology was chosen precisely because it is difficult to advocate for the concept of "order" in alignment when the music notation itself is so fluid and there is often no precise way to align voices. With this constraint removed by the use of OT rather than DTW, it is therefore possible that the methodology could match the first element in one setting with the last element in another, although this is unlikely due to the severe penalty this would incur. What is more common is that elements within the same area will be matched together, giving a more general sense of musical style than the exact matching process of DTW.

For example, one of the matches between *Sancte Germane* and *O Sancte Germane* as detected by the GNN–OT methodology can be seen in Fig. 7.23. All these notes appear in both settings, but not necessarily in this order, and they are distributed across all three voices rather than together in one line as presented here. Some matches may have been selected from successive places in the same voices, as would be expected from a classical matching methodology, but others may have been selected from further away at a higher cost, balancing a closer matching over correct temporal ordering. This is not to say that a set of notes will match regardless of order (which would render the comparisons meaningless) but that a note's order and position within the musical texture is already encoded within its vector representation. It is the vector representation that creates the match, of which order is a component in both the positional embedding and neighbourhood of each notational element. However, this positional data is currently irretrievable from the vector. Future work may make it possible to encode the exact positional data along with the the vector, so that the positional information that contributes to the match can be examined more closely. It is therefore impossible to look at the match presented in Fig. 7.23 and match it directly to any sequence as seen in the settings. What is being matched instead is the overall style: in this case a lot of Cs, a lot of As, and a lot of Gs next to As.

The clearest result in the GNN–OT methodology is the detection of overall musical style within genres. Fig. 7.24 visually presents the proportion of links between genres in all 1,069 results obtained by the GNN–OT methodology. In all genres except organa a3/4 and substitute clausulae, the largest proportion of all links were between settings of

Chord diagram showing genre links obtained by the GNN–OT methodology

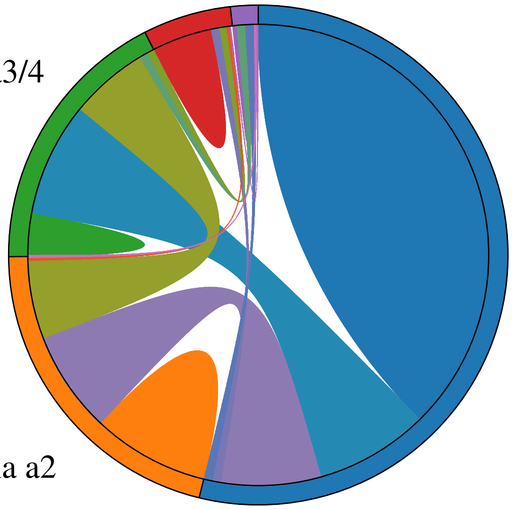

Fig. 7.24

the same style. It is of course not feasible that all of these connections can be interpreted as musical reuse, and most must be seen as false positives in this context. However, it is important to remark that these connections are not made randomly as one might expect if the methodology was unable to distinguish between genres: a conductus is most likely to be linked to another conductus, a two-voice organum to another two-voice organum. What we can infer from this is that the methodology is, at the very least, detecting something of the musical style of each genre. Features are successfully being extracted from conductus, two-voice organa, and motet that inform the methodology about the musical style being employed.

This may be due largely to the notational style. Within the results, it was found that motets were often linked most strongly by their final written passages, where the *cum littera* notation of the motetus ends and the regular *sine littera* ligature patterns of the tenor begin. Similarly, no other genre contains the same mellifluous notation as organum purum. Conducti move in a very distinctive note-to-note style, matched only by the intense discanting of the three- and four-voiced organa, and it is likely not chance, therefore, that the next largest set of links to and from conductus — aside from conductus itself — is to organa a3/4.

Nonetheless, it may be possible to read a little more into these results. A link between two settings is rarely detected by the HMM throughout the entire alignment. Signals such as line • in Fig. 7.21, showing the close alignment between two transmissions of *Sed sic eum* (M5, $\mathbf{W_1}$ f.27v–28r and $\mathbf{W_2}$ f.66v–67r) are in fact extremely infrequent. More commonly, the HMM in both the baseline and GNN–OT methodologies detects similarity for only part of the alignment: this is an intended behaviour, designed to capture short passages of musical reuse. Some of these matches are obvious false positives, especially links at the beginning and end where a sudden dip is caused by identical and likely incidental accidental and clef configurations. An example of such a false positive can be found in the • line, comparing the two-voice organum *Adiuua me domine* (M3) in $\mathbf{W_1}$ f.26v–27v with the conductus *Dic christi ueritas* in $\mathbf{W_2}$ f.33r–34v: although there are some dips in the signal, none are particularly low, and it is only the very end of the signal that is detected as perhaps being evidence for similarity. A signal such as this cannot be regarded, by itself, as musical reuse.

However, more difficult to discount are signals which clearly show some kind of pattern but cannot be explained by any typical musicological reasoning. Examples of this kind can be found in signals such as the • line, which compares the conducti *Puer nobis est natus* ($\mathbf{W_1}$ f.152r–154r) and *Ego reus confiteor* ($\mathbf{W_1}$ f.147v–148v). The signal contains a short yet undeniable dip roughly three quarters through its progress, indicating a region of similarity, but this cannot be corroborated with any clear motivation in the notation. A larger scale signal can be seen in the • line, which compares the two-voice organa *Dulce lignum* (M22, $\mathbf{W_1}$ f.34r–34v) and *Letabitur iustus* (M49, $\mathbf{W_1}$ f.45v–46r). Neither are listed as possessing an interrelationship between each other by Tischler (1988, pp.67, 74), but the plot begins with a clear low signal, a high peak, and then another low signal much like another manifestation of one of Smith's (1964) repertorial diagrams.

In none of these cases can a clear reason be found by manually scrutinising the notation and, much like with the melodic formulae critiqued in "2.9. Melodic formulae", confirmation bias can be a clear influence: a belief that there may in fact be some similarity can cause the mind to see new patterns in the notation, patterns that may in fact be insignificant and not contributing to the result obtained. However, it can be concluded that this methodology is capable of replicating many detections of musical reuse which have already been gained purely by manual musicological means. If a similarly convincing signal is generated for a concordance that is not immediately apparent in notation, questions must therefore be raised as to what is being detected in this case. The methodology's use of the entire corpus at once may be generating results not simply due to what is present but also perhaps due to what is missing: if two settings conspicuously omit the same figure in the same way, that can be considered just as strong a result in this methodology as a particularly conspicuous figure that is present in both settings. Many of these results may therefore contain low cosine difference due to what is not present rather than what is present, and this kind of result is much more difficult to verify.

7.8. Interpretation

One conclusion that can be made is that it is extremely difficult to interpret these results. The new GNN–OT methodology in particular was designed to extract patterns of musical reuse that may not be immediately visible, and highlight instead where the examined settings may play with similar textures. If these possible relationships were immediately apparent without this methodology, then they doubtless would have already been detected. Although many results detected by both baseline *Word2Vec*–DTW and GNN–OT methodologies can be corroborated by sources such as Smith's (1964) repertorial diagrams and Tischler's (1988) comparative edition, many results cannot. There are two ways in which this could be interpreted. The first would be to critique the results found here and argue that the results that are supported by extant literature were found only by chance or a partially operational methodology. The second interpretation would be to accept that the methodology does require refinement, but also argue that the results gained show that the central ideas are sound and may in fact point towards musical reuse at a deeper level than can be seen by point-to-point comparison of notation; further work is required to analyse these matches and trace the exact process that explains the result.

It is this second interpretation that is taken here: the results do show patterns, and although these patterns are not exact, they do much better than chance and provide an analyst with much more data than a manual process. Rather than simply analysing musical reuse as a surface process of copying, computational methodologies such as this one consider the whole corpus when analysing a section, and may provide insights into deeper structure than surface-level figurations; it could be likened to an X-ray that allows us to see the structure within.

However, it is difficult then to take these results, look back at the notation, and see how they relate. This is because the GNN–OT methodology's results do not simply map onto what can be seen in the notation, but rather how areas of notation are viewed within the context of the entire dataset. This is compounded by the fact that the GNN–OT results have no concept of exact sequence: notational elements can be drawn together from the whole setting in order to create a match, and it is currently difficult to control for this as each notational element is represented in an 18-dimensional latent space where only two dimensions control for time (16 embedding dimensions + 2 time dimensions).

We cannot therefore see a result that contains, for example, G3 and A3, and expect those two notes to appear in sequence within both settings. They may appear within the same locality in the setting, but they may not be directly linked. At the same time, they may appear only due to something else that is missing: there may be other elements that are more likely to appear within that setting but their absence in a match is conspicuous. The results therefore do not distinguish between elements that are notable due to their high probability and elements that are notable due to being unlikely. Further work is needed to link the matches obtained by the GNN–OT methodology back into the sources themselves. This would allow an analyst to observe not only the match between, for example, *Sancte Germane* and *O Sancte Germane* as in Fig. 7.23, but also where these notational elements have come from in each match, and thus whether the result is something that could be feasibly termed musical reuse or not.

It therefore becomes difficult to verify what we cannot see directly. We are accustomed to being able to immediately comprehend the reason exactly why a result has been generated, and so it is therefore difficult to equate the more intriguing results with tangible phenomena that we can see in the repertory. This is perhaps because these results demonstrate an aspect of the repertory that cannot be seen with the naked eye. However, this is exactly what this study has been working towards extracting: patterns of musical reuse that go beyond our typical conception of exact copying, clausula substitution, and identifiable melodic formulae. Indeed, it could well be argued that the digital methodologies such as those used in this study would be useless if they only served to verify what we already knew; the tools that this study has created and are used here allow us to see a little further beyond the immediate locality of the notation and patterns that we can detect through more traditional concordance techniques.

It is crucial therefore, to acknowledge the effectiveness of the CANDR system and dataset as a reliable data source. It has been carefully constructed to extract insights that are not readily apparent from a small-scale or surface-level survey of the notation, yielding results that not only affirm existing hypotheses through expected associations, but also prompt new conjectures. These findings extend beyond mere validation of existing theories. Instead, they encourage fresh new perspectives, ones that may evade conventional analysis; they are an indication of the approach's novelty and efficacy. As such, the emerging hypotheses resulting from these new methodologies open new avenues of inquiry, and offer insights that warrant further refinement and exploration. While the results presented here have revealed promising directions for further investigation, each of these findings demand more in depth scrutiny as in many cases the underlying reasons why they may be linked together are unknown. This underscores the fact that, while the obtained results hold inherent value, they represent only a preliminary step towards a more comprehensive understanding.

In summary, the results highlight the transformative potential of the alternative analytical lens demonstrated here, unveiling the possibility of some concealed connections between settings within the ND repertory, connections beyond what might typically be termed musical reuse. CANDR proves not only to be an adept tool for dataset creation but also a catalyst for further investigation. The implications of this signify not an endpoint, but a beginning to further study and deeper explorations of hitherto unknown connections within the ND repertory.

Chapter 8
Conclusion

> Philosophy is perfectly right in saying that life must be understood backward. But then one forgets the other clause — that it must be lived forward (Kierkegaard 1978, p.12).

It is abundantly clear that musical reuse within the ND repertory cuts far deeper into the creation, transmission, and preservation of ND polyphony than currently understood, and the issues that musical reuse highlights shape nearly every aspect of the study of thirteenth-century music. This study has brought the problem of how we might conceive of musical reuse more broadly within the ND repertory to the fore for the first time, confronting the influence of historiographical concepts on the research context. The clausula is considered both as an emblematic medieval fragment and as a force for the construction of historical narrative in modern-day study. Through the application of novel digital methods, I have critiqued interpretations that categorise musical reuse as a binary distinction, and pointed instead towards an alternate, looser, or perhaps intangible form of musical reuse within the ND repertory. The possible links found here are not explicitly formed by the copying and substitution of whole passages — such as is done in clausula substitutions and some indirect concordances — or the construction of polyphony through a patchwork of identifiable melodic formulae, but are demonstrated more generally through deeper signs of a common culture of musical knowledge shared by those involved with the repertory's transmission.

It has been 127 years since the introduction of the clausula-concept into the study of the ND repertory as part of Meyer's *Der Ursprung des Motetts* (1898), and 115 years since the first and yet unsurpassed survey of the repertory in Ludwig's *Repertorium* (1910). Early works such as these have been used, all too uncritically, as the foundational building blocks of our conception of the repertory. Each subsequent work has taken mostly as read the conclusions of its forbears, and built upon them to reach understandings that are a little further developed each time, in a reification of the popular idiom "standing on the shoulders of giants".[1] In attributing the creation of motet solely to a process of prosulating substitute clausulae, early scholars such as Meyer and Ludwig created an aetiology for the motet based on the wholesale

[1] The origin of the phrase is not, as commonly believed, Isaac Newton, but is in fact medieval. It can be traced at the very least back to an unsubstantiated quote by Bernard of Chartres in John of Salisbury's *Metalogicon* (McGarry 1971, p.167), but may also relate to the mythological characters of Orion and Cedalion.

substitution of like-for-like discant within the ND repertory, a narrative that viewed the motet as the end goal for musical reuse, and the clausula as the sole vehicle for accomplishing this. Although rightly now viewed with a high degree of scepticism, such simplistic narratives have created a culture of scholarship where the usefulness of terms such as "clausula" are maximised for our modern sensibilities, which are predominantly linear and causal.

However, the study of ND polyphony relies almost entirely on these second hand accounts: not just secondary literature, but contemporary accounts and music sources are also writings on a practice that, by the time of setting down, had already become a recollection. Careful study of both modern and medieval historiographic biases is only recently becoming an integral aspect of understanding and interpreting the ND repertory: appraisal of early scholarship in the field and the proclivity towards motet can be seen, respectively, in Busse Berger (2005, pp.9–44) and Beiche (2004), and a critique of the biases of Anonymous IV in van der Werf (1992) and Wegman (2015).

In addition to our own biases, the source texts themselves have their own biases. The ordering of the sources, how they are written, what they include (or, perhaps more importantly, what they leave out) are also all key indicators of the biases present in the sources themselves. For example, what was the agenda of the scribe of **F** in creating such a monumental manuscript? Why create a Magnus liber in the first place? Just like secondary sources, we cannot consider the primary musical notations as transparent communicators of an absolute music. They are snapshots of a living, breathing music culture, where the scribes are not only providing musical notation but also forging in parchment and ink their way of seeing (or indeed hearing) the music. We must consider not only what was written down but also what was perhaps left out, and the existence of musical reuse as an object beyond or deeper than what we can feasibly observe in the notation.

This study has aimed to deconstruct the idea of equating the concept of musical reuse with any specific terminology or terminologies, and put forward instead the idea of musical reuse as a broader, shared, and irreducible context in which thirteenth-century polyphony sits, a context that cannot be categorised into a series of independent phenomena such as clausulae substitutions, indirect concordances, or melodic formulae, but which must instead be expressed as part of an overlapping continuum of musical reuse. This is the natural result of a process of musicians 'creating melodies like the ones they sang every day' (Allison 2020, p.46).

Can we come up with definitions of terms such as "clausula" that we can be sure are equivalent to medieval usage for the ND repertory? Moreover, can we decide on definitions that are consistent within modern literature? Based on the discussion in "Chapter 2: Historiography of musical reuse: the clausula and its context", the answer to both of these questions is, unequivocally, no. If the works of post-structuralist philosophers such as Foucault account for anything, then it is that there is no objective view of the past (Foucault 1972). All that truly can currently be ascertained with any certainty are the basics, such as those outlined in "1.4. Another look". Following from this thinking, the aims in this dissertation therefore 'change from documenting the past to deconstructing the existing scholarly narratives' (Kurkela and Mantere 2015, p.3). Whether that past is 5, 50, or 500 years gone makes little difference: it must still be reconstructed and re-understood through a reflection of our own biases. However, as Wegman (2012, p.43) observes:

> Like Narcissus, or so critics remind us, we have gazed into a fountain, and have become enamored of the image reflected in its surface. The fountain, one might say, is the totality of the available historical evidence, and the images it returns is the product of our historical vision. We have wanted that image to be real, objective, autonomous, authentic, other. Like Narcissus, however, we have been frustrated in our attempts to capture the image — that is, to demonstrate its objective reality.

In short, 'we have been duped [...] we have made fools of ourselves' (Wegman 2012, p.44). Regardless of the mode of inquiry, any "historical truth" that we may purport to unearth is inherently subjective, based on our own understanding and discursive relationship with alternate understandings from the past, both near and far. There is therefore not a single and objective concept of musical reuse that we can alight upon, as each of our understandings are viewed through our own biases and methodological approaches.

The computational and numerical methodologies used in this study, which could be thought of as attempts to objectify the research, in fact remain in many ways subjective interpretation, and are only part of a myriad of possibilities for investigating the issue: it is important to account for the fact that I selected these methodologies simply because I believed that they would yield results useful to my argument. As has been made clear in my discussion, each part of the methodology required a series of decisions to be made, and parameters to be set. Granted, some of these have objective, "best" answers, but others were chosen experimentally, and some had to be chosen arbitrarily. These results are key to my understanding, they may in future form part of another person's understanding, but are not historical fact. The best I can hope for, then, is to answer the question: "What is this study's understanding of musical reuse?" This, too, must be incomplete as there are many related aspects that I have elected to omit: the relationship between clausula and motet, alternative etymologies, an explicit aspect of orality, and where copula fits into this. These have all been left out in favour of a historiographical approach that does not necessarily tell us anything more about historical evidence per se, but rather the issues surrounding the understanding of currently available evidence and argument. This is the basis upon which my understanding lies.

At the same time, this dissertation has discussed similar issues in the search for objective "truth", through the stumbling blocks on the path to producing a satisfactory digital editing of ND notation via the representation of the semantic meaning of the notation in the computer. The advent of the computer has not solved any of the traditional issues of music editing, and Bent's (1994, pp.391–392) argument still holds that:

> The original notation […] is the only authority to which new editions and performances can turn for the notes and rhythms. The modern transcription, in providing a sound map of a piece that is more prescriptive for modern users, may sacrifice some of the dimensions present in the original, as any translation represents loss.

An edition, in transcription or translation, may represent the notation from which it is drawn, but cannot and must not supersede the source itself: there is both the edition as medium for the musical "message", and the edition as musical simulacrum.

It could be argued that the application of the digital humanities may even go so far as to exacerbate the problem. For one, the addition of the computer adds yet another layer of fraught translation to the process of edition. More significantly, however, the added possibilities for instant conversion, analysis, and comparison in the computer aggravate the complexity of the situation by expanding the horizons for which the digital representation is capable. Whereas the paper edition only had to deal with the resultant graphical domain of the final copy, a digital edition must also come to terms with any future conversions and interpretations: not just the graphical domain but also the logical, analytical, audial, and beyond. Without a clear picture of what the addition of the computer is going to be useful for, the digital turn in early music scholarship can very easily inherit anachronistic ideas and assumptions, both by a one-sided or out-of-date scholarly representation creating a sense of digital consensus where there is none, or perfectly reasonable assumptions in one repertory being unfairly imported wholesale into others.

Taking his inspiration from Douglas Adams, Dumitrescu (2013, p.213) argued that for early music editing 'we still struggle to decide precisely what it is we want or do not want from our machines […] do they give us answers […] or do they help us to find out just what the questions are?' Not defining which of the two we are searching for leads us simply to the saying "If you automate a mess, you get an automated mess".[2] I have

[2]This idiom has been attributed to countless individuals, but most commonly to Rod Michael and Michael Martin Hammer. The exact origin is unknown.

argued in "Chapter 3: Data models" that ND notation is particularly vulnerable to these concerns. Any edition of music faces a broad set of issues, but the digital editing of early polyphony faces particularly unique problems partly due to its historiographical position between monophonic plainchant on the one hand and mensural notation on the other. It has features of both — partly unmeasured like plainchant, polyphonic like later mensural music — but operates like neither. Whereas a sequence of monophony can be described by a series of events, and the rhythm of mensural notation largely by the editorialisation of its mostly static and rational rules, early polyphony such as ND notation cannot be adequately described in sequence, nor should its rhythm be "worked out" when there is little consensus as to how that rhythm truly works. Once again, there is little historical fact to be found here: perhaps there are a series of "house styles" that can be applied to particular sections of the repertory (Roesner 1981, p.375), but there are no hard and fast rules. This is despite the writings of theorists such as Anonymous IV who perhaps would have liked to paint a picture of the repertory as subject to perfectly rational rules, and even purum being able to be described 'exactly the same as in the other regular modes' (Yudkin 1985, p.77).

On the contrary, the typical data structures used for encoding music in the computer — such as tables and hierarchies — require fully prescriptive notation for polyphony, particularly when it comes to the two axes of pitch and rhythm. This is not something that can be agreed upon for ND notation, as more often than not the simultaneity of two voices is inferred editorially and solely by graphical alignment. In three- and four-voice music, these points of alignment are inferred on a per-voice basis — for example in hocket — and cannot be used to uniformly divide a data structure. Common music encoding schemata and analysis frameworks are not designed to support this ambiguity and often editorialise any ambiguous notation into more prescriptive forms for analysis. For example, Cumming et al (2018, p.493) admit that our current tools are insufficient in this regard:

> Although there are often good musicological reasons for ambiguity, it can cause serious problems for many systematic analysis, search, display, or feature extraction systems, which may use improper defaults or not work at all when faced with certain kinds of ambiguous data.

It would be possible to editorialise ND polyphony into prescriptive and aligned rhythm to fit currently available schemata and frameworks, but altering our evidence simply to fit our current tools seems to me to be deeply unscientific and backwards. Forcing the evidence into the restrictive moulds used by inadequate tools may yield results easily, but these results would tell us more about the tools than the evidence.[3] Rather, the nature and significance of the evidence should be considered first, and the tools designed around the problem at hand. This is the undercurrent of the approach that I have articulated throughout this discussion.

I have argued that a study such as this one that is dealing with ambiguous graphical notation should therefore begin in the same graphical domain and do as much as possible with a graphical encoding before any attempts to translate the model into a logical domain because, as Bent (1994, p.392) reflects above, 'any translation represents loss'. The later we make this translation and the less that requires translation, the better. Moreover, any translation not only loses information but adds a new, subjective layer — this could be described as the will or bias of the editor — and, as above, it is impossible for this layer to be grounded in any "historical fact": it reflects just as much an image of the editor and their culture as the edited artefacts themselves. Digital analysis cannot avoid this issue, but this study has carefully considered the domains in which it has operated and made plain its exact transformations between the graphical, logical, and analytical.

This issue extends not just to encoding but also to the methodologies chosen for analysis. Any number of analysis techniques could have been chosen as part of this study, and my reliance on digital methods does not in any way absolve me from any of these epistemological concerns. What it has enabled, however, is the foregrounding of these same issues. A human editor or analyst thinks very little of a small exception, or even of creating a new model just to describe one slice of the evidence. Indeed, the ND repertory is full of these small exceptions. The introduction of the computer into the mix forces us to consider these exceptions on the same level as the rest of the evidence: whereas a human will often happily bend or break their model to accommodate an exception, relegating the exception itself to footnote status, the systematic and rigid manner in which music must be described in the computer does not allow us to avert our eyes for a few exceptions. If some encoded notation does not quite fit into our computer model, then to the computer it might as well be completely alien: it either fits or it does not.[4]

Rather than forcing all the music into a model that does not fit every single possibility within the notation, I have considered problematic passages for encoding as the guide for how the entire repertory should be encoded, rather than a preferred rhythmical or rational subset. One common denominator, beyond the already apparent notes, clefs, and ligature, are the graphical links left-to-right from one note to the next, and the ambiguous yet necessary editorial alignments of voice synchronicity. As a

[3] Another proverbial saying, frequently attributed to Abraham Maslow: "If the only tool you have is a hammer, it is tempting to treat everything as if it were a nail".

[4] Recent advances in Deep learning and AI are not exceptions to this, it is simply that the models are larger and more nuanced. In the machine learning space, this very humanistic breaking of rules when it suits can be placed under the broader concept of "concept drift", and it is still a largely unsolved problem for machine learning (Lu et al 2019).

result, I have developed a graph-based encoding format that can accurately encode these two features without over-editorialisation or condensing ambiguous notations into prescriptive rhythms.

The rejection of common encoding schemata in this dissertation has also led me to reject common analysis techniques. In this respect this study has been unique in two ways. First the size of its dataset far outstrips that of Falck (1990) or Flowers (2013) who only considered the same small sample of conductus, or Erickson (1970) whose tabular dataset considered only organum duplum. Secondly, the unsupervised and semi-supervised methodology that I have used has not been attempting to prove a theory by the analysis of fully labelled data, but has taken mostly unlabelled data and formulated hypotheses based upon the clusters created. Whereas manual analysis often takes a broad or perhaps presumptive view when it comes to the question of whether two pieces are "the same" or "similar", the quantitative analysis considered here has forced me to define concretely what is meant by these terms, and therefore guided me towards an analysis with a high degree of rigour.

Taking my cue from modern machine learning methodologies and Big Data, I have endeavoured to analyse not a sample but the entire dataset. Feature extraction is typically aligned along the two axes of pitch and rhythm, but as the encoding methodology taken here does not use these axes, typical analysis software could not be used, and typical features could not be extracted. I developed a feature extraction methodology that takes the graph-based encoding format used in this study to create rich node embeddings, where not only the type of notational element but also its surrounding context alter the element's expression in vector format. I also developed a new methodology for aligning areas of similarity together, not using n-grams (due to their difficulty to reconcile with polyphony) but using time series alignment techniques on the cosine distance between pairs of polyphonic settings.

This work is manifested in the *Clausula Archive of the Notre Dame Repertory* (CANDR), a system that I developed from scratch as part of this project.[5] It consists of two main parts. The first, described in "Chapter 5: Online", is a public-facing web application that allows users to view the facsimiles of the notational sources as well as linked catalogues of the various aspects (e.g. by folio or by setting), browse the dataset created as part of this project, view and edit tracings of notational elements on the staves and systems in an intuitive and graphical user interface backed by a custom OMR system, and export automatically generated symbolic notation. This dataset and repository alone will be of great use and impact to scholars working in this field: never before have both the graphical domain of the source facsimiles and a logical domain of

[5] <https://www.candr.org.uk/>.

tracings and transcriptions been so closely linked. For the first time, this web application allows scholars, students, or anyone curious about the repertory to explore ND polyphony from the broadest level of the sources and facsimiles down to the minutiae in individual notes.

The second part of CANDR, described in "Chapter 6: Offline", is a *Python* analysis toolkit, based on and interoperable with common scientific computing libraries *NumPy*, *SciPy*, *Scikit-learn*, *Pandas*, and *PyTorch*, that implements the analysis methodologies described in this dissertation. The toolkit consists of a framework for creating an analysis in the form of a Big Data pipeline, by isolating each stage within a single process and transparently streaming data between each, allowing a user to create any number of flexible and exploratory analyses. Analysis stages are created for fetching data from the CANDR database, filtering, loading, saving, parsing, analysing, and exporting results. In this study's analysis, results have been generated by the application of the programming toolkit to the dataset that I created within the web application.

With regard to the results themselves, some of what is uncovered is (reassuringly) hardly surprising: settings that we understand as "the same" between two sources display a high degree of similarity, substitute clausulae are similar to settings where those clausulae are substituted. A computational analysis was not needed for such conclusions, as the similarities are immediately apparent in the notation, and many are already catalogued in catalogues raisonnés such as van der Werf (1989), as well as the critical apparatus of editions such as Tischler (1988) and Payne (1996). Rather, the more interesting portion of the results are the subset of those found in the GNN–OT methodology but not the baseline methodology (i.e. the furthest right segment of the Venn diagram in Fig. 7.17): these are results that are not replications of already-known, surface-level musical reuse, but yet demonstrate signals in the GNN–OT methodology consistent with musical reuse. These signals may in fact be evidence of connections between settings of music at a level not immediately visible to the naked eye. The similarities this methodology suggests between supposedly dissimilar settings are not apparent in a comparison of notation, demonstrating that there may be musical reuse within ND polyphony beyond visible similarities. In particular, the GNN–OT methodology developed here may be extracting not only similarities between passages of ND polyphony that are known or easily observed (the intersection with the baseline methodology as well as those that the baseline methodology missed), but also similarities that are hidden within the music, or existent at a deeper layer than surface-level figuration. To answer Dumitrescu's question cited above (2013, p.213), this machine provides many more questions than answers.

Such a result should be expected, although it is a little unnerving to accept results that we cannot verify by close reading. In the study of music, or perhaps more commonly within the humanities, any method or analysis that we may use is secondary to human intuition: hypotheses are generated by a scholar's connection with the material and a tentative understanding that may arise from it. This can be formulated as a testable hypothesis, which can then be confirmed or denied by the application of the scientific method. However, especially in the more abstract sciences, instruments are used to create, manipulate, and test the material. It is not uncommon for the object in question never to be observed directly but the effects discovered entirely through numeric or projective means. For example, fundamental particles are not directly observed by physicists but the results of their decay can be observed by detectors, similarly exoplanets are not observed by astronomers but their size and position can be inferred by their gravitational and dimming effects on nearby stars. This raises complex epistemological questions on the nature of observation (Pacherie 1995), yet it remains that instruments such as microscopes, telescopes, and particle accelerators are considered valid tools despite the fact that observations made by them cannot be verified by the naked eye.

The results in this study follow that same, indirect method. The corpus of music created as part of this project consists of over half a million items distributed over nearly 10,000 staves. The analysis process has considered all items and all staves at once. Therefore, when a result has been generated, such as "the pairwise cosine similarity between the aligned node embeddings is much lower here than expected in the corpus as a whole", there are two main reasons why it cannot be verified manually. First, the OT methodology used here does not map to sequence nor to any discernible notational object within the music: it is simply a vector representation of the notation generated using an unsupervised methodology, and some of the matching may be due to the fact that other elements may not be present. Secondly, a match between two settings may include notational elements that are not in close proximity to each other, but the close similarity between elements may lead to a lower overall score. Both the DTW and OT methodologies are attempting to minimise the overall distance between signals, and therefore the length of each signal may affect the alignment. It is best then not to conceive of the matches in the results as mapping directly to specific and intended patterns of musical reuse. Rather, the matches are proxies for certain vector effects within the latent space mapped out by the corpus. The only two ways, then, that the results obtained in this methodology can become more convincing are by linking these matches back into their musical and notational context in the sources, and by further evaluation and verification of the methodology on already known effects. Both of these are possible, but would require more extended study to prepare the necessary techniques and data to incorporate them into the analysis.

The main limitation of this study is precisely this same issue of verification. The methods by which, and reasons for, many of the results generated are as yet not situated in a purely musicological and theoretical context. For example, the output results of potential new evidence of musical reuse, with similar patterns to outputs of extant examples, although remarkably similar to signals that replicate known similarities within the corpus, are difficult to situate with a clear musicological reasoning. If we accept that the methodology does provide meaningful results due to its replication of known musical reuse, then such signals need to be seriously considered, but at the same time the results cannot be attributed to anything that we already see in the repertory, beyond a general impression that may be gained from a wide-ranging knowledge of the notation and polyphony. What this study has not provided are tools at a higher level beyond the technical that would allow us to investigate further why such distributions and connections emerge from the data using this methodology.

Nonetheless, the results demonstrate that the issues surrounding musical reuse within the ND repertory are not as simple and categorical as we might like to conceive. For over a century, we have known that whole sections of polyphony — particularly discant — were swapped between settings, such that a setting transmitted in one source may be almost identical to that transmitted in another, with the exception of certain passages that transmit radically different polyphony. However, there have always been hints that such verbatim substitution is only the tip of the iceberg in the context of musical reuse within the repertory: e.g. the suggestion that some substitute clausulae do not fit very well (Frank Ll. Harrison 1958, p.123), theorised links and indirect concordances between conductus and clausula and unrelated organa (Bukofzer 1953), melodic formulae (see "2.9. Melodic formulae"), evidence of oral processes within the repertory (Busse Berger 2005), as well as the influence of vernacular song on liturgical polyphony (Bradley 2013). The results presented here are part of this same trend of analysing exceptions, of problematising our linear and causal historical narratives not only because they cannot explain the whole picture, but also because they may lead to false or all-too-tidy results.

This study has extended these ideas from symptoms of a complex picture of independent musical reuse processes into a broader framework that re-contextualises the relative importance of manifestations of musical reuse such as clausulae and melodic formulae within the possibility of a larger and perhaps unwritten culture of musical reuse. The substitutions apparent between clausulae are doubtless some of the most apparent forms of musical reuse within the repertory, but they do not comprise its entirety. More complex patterns of musical reuse may have been part and parcel of oral processes, a shared pedagogy, or a deeper understanding of formulaic reuse that was present in the twelfth and thirteenth centuries that has since been lost. Whatever the case may be, this dissertation has begun to unearth its remaining evidence within the notational sources of the thirteenth century.

The result of this is that this study opens up countless avenues for further research. I believe that they fall into four broad categories: historiographical, methodological, evidential, and theoretical. First, in the historiographical aspect, more research should be done to critique the terms and terminologies that are used within the study of ND polyphony and more broadly within medieval music. We should ascertain exactly how these terms are understood and used both in medieval and modern usage, the etymologies that have been created for them, and the contexts in which they have been invoked to conjure up particular connotations. This dissertation has begun this process for the term "clausula" and I have hinted at this same problem for "motet" in Stutter (2023), but there are other terms that require the same analysis, such as "Magnus liber", "copula", and even the distinction between what constitutes "monophony" and "polyphony".

Second, in the methodological aspect, the techniques created as part of this study are almost entirely novel, and there are no prior examples of this analysis methodology used in digital musicology as a whole. As a result, the methodology itself requires further evaluation and testing on this repertory as well as others to better understand its significance and operation. Moreover, more powerful and user-friendly tools should be created to interact with this methodology: it is currently a highly technical and opaque technique, which limits its usefulness for less technically-minded musicologists. Work should therefore be done to make the methodology more transparent, not merely to be able to inspect why certain results are generated and to weed out the possibility of false positives, but also to improve the applicability of the methodology by making its processes more easily presentable and explainable.

Third, in the evidential aspect, the CANDR dataset as it currently stands does not cover even half of all the polyphony available in digital facsimile. Indeed, the choice of sources $\mathbf{W_1}$ and $\mathbf{W_2}$ in this study, although warranted in this case, has limited the possibilities for organa–clausulae overlap: these two sources have different foci, and although both contain organa, motet, and conductus, $\mathbf{W_1}$ is more focused on substitute clausulae and $\mathbf{W_2}$ on vernacular motet. The main candidate for further research therefore is **F** which contains much of the material available in both $\mathbf{W_1}$ and $\mathbf{W_2}$, but there are many other peripheral or fragmentary sources that could also contribute valid material to the database. Granted, many pieces are *unicae*, but if my hypothesis of further musical reuse has merit then it may be possible that these *unicae* have concordances with other settings.

Finally, in the theoretical aspect, it may be noted throughout this dissertation that although I have made advancements in evidence and results, there is not yet a robust theoretical and musicological framework in which to use them. Put another way, some of the results presented as part of this study can be seen as evidence for questions not yet asked. There is a lacuna of academic scholarship concerning deeper and more nuanced repertorial structures within and beyond ND polyphony, and further research must be done to theorise the possibility of broad musical reuse on a musicological and cultural level. It is only once a framework for how it operates is theorised that this and other new methodologies can be used more incisively to test and to provide evidence for these theories.

This dissertation has taken a fresh look at the ND repertory, its issues, problems, and potential solutions. I have reappraised the historiographical impact of musical reuse in the thirteenth century by advocating for musical reuse to be considered holistically rather than as a set of separate phenomena. To achieve this, I have developed novel Big Data digital humanities methodologies and implemented them in a large-scale web application and analysis framework. Given all of this, we can finally return to my original question: How should musical reuse be defined in the ND repertory?

Musical reuse is a concept tightly bound up with the ND repertory's important historiographical position in Western art music. For better or worse, the historiographical position of the ND repertory has been in part elevated by music historians due to its musical reuse, rather than in spite of it. This is especially the case for the clausula, which has become symbolic of musical reuse in the thirteenth century, and a vital narrative stepping stone for constructing the history of the motet. However, musical reuse can be seen almost everywhere in the ND repertory, and although large scale reuse is the most obvious, small scale reuse can be seen in melodic formulae and common patterns. Modern study has created categories for musical reuse, given them names, and assigned criteria to each. However, I believe that it is unlikely that thirteenth-century musicians thought in such categorical terms, and there is scant evidence for our modern taxonomies in medieval contexts. Moreover, it is not possible to reconstruct these medieval concepts in modern study, as we are searching for a sense of "historical fact" that never existed.

Large-scale analysis does not bear out these categories either, and rather than clear separations between "similarity" vs "difference" or a series of clearly-distinguishable types of musical reuse, measuring reuse yields results more continuous and uniform than we would expect. In only the most obvious cases do the plots show delineated cases of reuse where an obvious change in signal occurs and the two settings engage in the same music in the same way. Instead, although we see areas of high similarity and low similarity, we see many more shades in between, and there are infinite gradations of similarity. Discussing similarity between settings of ND polyphony using categorical language requires drawing an arbitrary line at one or more of these gradations. As such we must more clearly define what we mean when we discuss this more nuanced idea of reuse or, better, develop methodologies for measuring it. A future view of musical reuse therefore may dispense even with "reuse" as a simple yes or no category: the question is not "is there musical reuse?" but rather, "how much?"

References

Achimugu, Philip, Ali Selamat, Roliana Ibrahim, and Mohd Naz'ri Mahrin. "A systematic literature review of software requirements prioritization research." *Information and Software Technology* 56 (2014) : 568–585 <https://doi.org/10.1016/j.infsof.2014.02.001>.

Allison, Chloë Nicola. *The Creation of Parisian Organum Purum: Office Organa Dupla in the MLO Sources*. PhD diss. Cambridge: University of Cambridge, 2020 <https://www.repository.cam.ac.uk/handle/1810/331048>.

Antoniou, Grigoris and Frank van Harmelen. *A Semantic Web Primer*. Cambridge, MA: The MIT Press, 2004.

Apel, Willi. *The Notation of Polyphonic Music 900–1600*. 1941. 4th ed. Cambridge, MA: The Mediaeval Academy of America, 1949.

Babbitt, Milton. "The Use of Computers in Musicological Research." *Perspectives of New Music* 3 (2) (1965) : 74–83 <https://doi.org/10.2307/832505>.

Bainbridge, David and Tim Bell. "The Challenge of Optical Music Recognition." *Computers and the Humanities* 35 (2) (2001) : 95–121 <https://doi.org/10.1023/A:1002485918032>.

Baltzer, Rebecca A. *Notation, Rhythm, and Style in the Two-voice Notre Dame Clausula*. PhD diss. Boston, MA: Boston University, 1974.

———. *Le Magnus Liber Organi de Notre-Dame de Paris. Volume V: Les Clausules a Deux Voix du Manuscrit de Florence, Biblioteca Medicea-Laurenziana, Pluteus 29.1, Fascicule V*. Monaco: Éditions de L'Oiseau-Lyre, 1995.

Baratè, Adriano, Goffredo Haus, and Luca A. Ludovico. "A critical review of the IEEE 1599 standard." *Computer Standards & Interfaces* 46 (2016) : 46–51 <https://doi.org/10.1016/j.csi.2016.02.001>.

———. "State of the Art and Perspectives in Multi-Layer Formats for Music Representation." *2019 International Workshop on Multilayer Music Representation and Processing (MMRP)* (2019) : 27–34 <https://doi.org/10.1109/mmrp.2019.00014>.

Baró, Arnau, Pau Riba, Jorge Calvo-Zaragoza, and Alicia Fornés. "From Optical Music Recognition to Handwritten Music Recognition: A baseline." *Pattern Recognition Letters* 123 (2019) : 1–8 <https://doi.org/10.1016/j.patrec.2019.02.029>.

Baró, Arnau, Pau Riba, and Alicia Fornés. "Musigraph: Optical Music Recognition Through Object Detection and Graph Neural Network." *Proceedings of the 18th International Conference on Frontiers in Handwriting Recognition (ICFHR)* (2022) : 171–184 <https://doi.org/10.1007/978-3-031-21648-0_12>.

Beiche, Michael. “Motette.” *Handwörterbuch Der Musikalischen Terminologie*. Ed. Eggebrecht, Hans Heinrich. 4. Stuttgart: Franz Steiner, 2004.

Bell, Nicolas. “Reviewed Work(s): Le Magnus Liber de Notre-Dame de Paris by Mark Everist.” *Early Music History* 25 (2006) : 309–323.

Bent, Margaret. “Editing Early Music: The Dilemma of Translation.” *Early Music* 22 (3) (1994) : 373–392 <https://doi.org/10.1093/earlyj/xxii.3.373>.

———. “Early Music Editing, Forty Years On: Principles, Techniques, and Future Directions.” *Early Music Editing: Principles, Historiography, Future Directions*. Eds. Dumitrescu, Theodor, Karl Kügle, and Marnix van Berchum. Turnhout: Brepols, 2013.

Berry, David M. and Anders Fagerjord. *Digital Humanities: Knowledge and Critique in a Digital Age*. Cambridge: Polity Press, 2017.

Bevilacqua, Gregorio, David Catalunya, and Nuria Torres. “The Production of Polyphonic Manuscripts in Thirteenth-Century Paris: New Evidence for Standardised Procedures.” *Early Music History* 37 (2018) : 91–139 <https://doi.org/10.1017/S0261127918000049>.

Biblioteca Medicea Laurenziana. *Scaffale Digitale: Plut.29.1*, n.d. <http://mss.bmlonline.it/s.aspx?Id=AWOHy_N-I1A4r7GxMB57&c=Antiphonarium#/book>.

Bohlman, Philip V. “Ontologies of Music.” *Rethinking Music*. Eds. Cook, Nicholas and Mark Everist. Oxford: Oxford University Press, 2001. 17–34.

Botsch, Michael-Felix. *Machine Learning Techniques for Time Series Classification*. Göttingen: Cuvillier Verlag, 2023.

Box, George E. P. and Norman R. Draper. *Empirical Model-Building and Response Surfaces*. New York: John Wiley & Sons, 1987.

Bradley, Catherine A. “Contrafacta and Transcribed Motets: Vernacular Influences on Latin Motets and Clausulae in the Florence Manuscript.” *Early Music History* 32 (2013) : 1–70 <https://doi.org/10.1017/S0261127913000016>.

———. “Re-workings and Chronological Dynamics in a Thirteenth-Century Latin Motet Family.” *The Journal of Musicology* 32 (2) (2015) : 153–197 <https://doi.org/10.1525/jm.2015.32.2.153>.

———. “Origins and Interactions: Clausula, Motet, Conductus.” *A Critical Companion to Medieval Motets*. Ed. Hartt, Jared C. Woodbridge: The Boydell Press, 2018a. 43–60 <https://doi.org/10.1017/9781787442542.004>.

———. *Polyphony in Medieval Paris: The Art of Composing with Plainchant*. Cambridge: Cambridge University Press, 2018b <https://doi.org/10.1017/9781108290456>.

———. “Choosing a Thirteenth-Century Motet Tenor: From the Magnus liber organi to Adam de la Halle." *Journal of the American Musicological Society* 72 (2) (2019) : 431–492 <https://doi.org/10.1525/jams.2019.72.2.431>.

Bradley, Catherine A., Christoph Wolff, Graham Dixon, James R. Anthony, Malcolm Boyd, Jerome Roche, Leeman L. Perkins, Ernest H. Sanders, Peter M. Lefferts, and Patrick Macey. "Motet." *Grove Music Online*. Oxford: Oxford University Press, 2019 <https://doi.org/10.1093/omo/9781561592630.013.90000369371>.

Bronson, Bertrand H. "Toward the Comparative Analysis of British-American Folk Tunes." *The Journal of American Folklore* 72 (284) (1959) : 165–191 <https://doi.org/10.2307/538481>.

Buisson, Morgan, Brian McFee, and Slim Essid. "Using Pairwise Link Prediction and Graph Attention Networks for Music Structure Analysis." *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)* (2024) <https://hal.science/hal-04665063v1>.

Bukofzer, Manfred. "Interrelations between Conductus and Clausula." *Annales Musicologiques* 1 (1953) : 65–103.

———. "Discantus." *Musik in Geschichte und Gegenwart*. Ed. Blume, Friedrich. 3. Basel: Bärenreiter-Verlag, 1954. 559–578.

Burgoyne, John Ashley, Ichiro Fujinaga, and J. Stephen Downie. "Music Information Retrieval." *A New Companion to Digital Humanities*. Eds. Schreibman, Susan, Ray Siemens, and John Unsworth. 1st ed. Oxford: John Wiley & Sons, 2016. 213–228 <https://doi.org/10.1002/9781118680605.ch15>.

Burkholder, J. Peter. *All Made of Tunes: Charles Ives and the Uses of Musical Borrowing*. New Haven, CT: Yale University Press, 1995.

Burkholder, J. Peter, Donald Jay Grout, and Claude V. Palisca. *A History of Western Music*. 7th ed. New York, NY: W. W. Norton & Company, 2006.

Burney, Charles. *A General History of Music*. 1789. Ed. Mercer, Frank. 2nd ed. Harcourt, Brace and Co.: New York, NY, 1935.

Busse Berger, Anna Maria. "Mnemotechnics and Notre Dame Polyphony." *The Journal of Musicology* 14 (3) (1996) : 263–298 <https://doi.org/10.2307/764059>.

———. "The evolution of rhythmic notation." *The Cambridge History of Western Music Theory*. Ed. Christensen, Thomas. Cambridge: Cambridge University Press, 2002a <https://doi.org/10.1017/chol9780521623711.022>.

———. "Review of *Perotinus Magnus* by Jürg Stenzl." *Plainsong & Medieval Music* 11 (1) (2002b) : 88–98 <https://doi.org/10.1017/S0961137102232064>.

———. *Medieval Music and the Art of Memory*. London: University of California Press, 2005 <https://doi.org/10.1525/california/9780520240285.001.0001>.

Büttner, Fred. *Das Klauselrepertoire der Handschrift Saint-Victor (Paris, BN, lat. 15139)*. Milella: Lecce, 2011.

Byrd, Donald A. *Music Notation by Computer*. PhD diss. Bloomington, IN: Indiana University, 1984.

Cage, John. *Notations*. New York, NY: Something Else, 1969.

Calvo-Zaragoza, Jorge, Jan Hajič Jr, and Alexander Pacha. "Understanding Optical Music Recognition." *ACM Computing Surveys* 53 (4) (2020) : 1–35 <https://doi.org/10.1145/3397499>.

Calvo-Zaragoza, Jorge and David Rizo. "End-to-End Neural Optical Music Recognition of Monophonic Scores." *Applied Sciences* 8 (2018) <https://doi.org/10.3390/app8040606>.

Calvo-Zaragoza, Jorge, Alejandro H. Toselli, and Enrique Vidal. "Hybrid hidden Markov models and artificial neural networks for handwritten music recognition in mensural notation." *Pattern Analysis and Applications* 22 (2019) : 1573–1584 <https://doi.org/10.1007/s10044-019-00807-1>.

Calvo-Zaragoza, Jorge, Gabriel Vigliensoni, and Ichiro Fujinaga. "A Machine Learning Framework for the Categorization of Elements in Images of Musical Documents." *Proceedings of the Third International Conference on Technologies for Music Notation and Representation (TENOR)* (2017a) : 17–24 <https://www.tenor-conference.org/proceedings/2017/02_Calvo-Zaragoza_tenor2017.pdf>.

———. "Pixelwise classification of music document analysis." *Proceedings of the Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA)* (2017b) <https://doi.org/10.1109/IPTA.2017.8310134>.

Carter, Bryan W. "Introduction to the Digital Humanities." *Digital Humanities: Current Perspective, Practices, and Research*. Cutting-edge Technologies in Higher Education. 1st ed. Emerald Group Publishing: Bingley, 2013. 1–10.

Castellanos, Francisco J., Gabriel Vigliensoni, Jorge Calvo-Zaragoza, and Ichiro Fujinaga. "Document Analysis of Music Score Images with Selectional Auto-Encoders." *Proceedings of the 19th International Conference on Music Information Retrieval (ISMIR)* (2018) : 256–263 <https://ismir2018.ismir.net/doc/pdfs/93_Paper.pdf>.

Cecire, Natalia. "When Digital Humanities Was in Vogue." *Journal of Digital Humanities* 1 (1) (2011) <http://journalofdigitalhumanities.org/1-1/when-digital-humanities-was-in-vogue-by-natalia-cecire/>.

Cerquiglini, Bernard. *In Praise of the Variant: A Critical History of Philology*. Trans. Betsy Wing. London: The Johns Hopkins University Press, 1999.

Chizat, Lenaïc. *Numerical methods for (unbalanced) optimal transport*, 2016 <https://github.com/lchizat/optimal-transport>.

Chuan, Ching-Hua, Kat Agres, and Dorien Herremans. "From context to concept: exploring semantic relationships in music with Word2Vec." *Neural Computing and*

Applications 32 (2020) : 1023–1036 <https://doi.org/10.1007/s00521-018-3923-1>.

Ciula, Arianna and Cristina Marras. "Exploring a Semiotic Conceptualisation of Modelling in Digital Humanities Practices." *Meanings & Co.: The Interdisciplinarity of Communication, Semiotics and Multimodality*. Eds. Olteanu, Alin, Andrew Stables, and Dumitru Borţun. Numanities – Arts and Humanities in Progress. New York, NY: Springer, 2018. 33–52 <https://doi.org/10.1007/978-3-319-91986-7_3>.

Codd, Edgar Frank. "A relational model of data for large shared data banks." *Communications of the ACM* 13 (6) (1970) : 377–387 <https://doi.org/10.1145/362384.362685>.

Cook, Nicholas. "Epistemologies of Music Theory." *The Cambridge History of Western Music Theory*. Ed. Christensen, Thomas. Cambridge: Cambridge University Press, 2002. 78–105 <https://doi.org/10.1017/chol9780521623711.005>.

———. "We Are All (Ethno)musicologists Now." *The New (Ethno)musicologies*. Ed. Stobart, Henry. Plymouth: The Scarecrow Press, 2008. 48–70.

Cornelissen, Bas, Willem Zuidema, and John Ashley Burgoyne. "Mode Classification and Natural Units in Plainchant." *Proceedings of the 21th International Conference on Music Information Retrieval (ISMIR)* (2020) <https://archives.ismir.net/ismir2020/paper/000232.pdf>.

Cosenza, Emanuele, Andrea Valenti, and Davide Bacciu. "Graph-based Polyphonic Multitrack Music Generation." *Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI)* (2023) <https://doi.org/10.24963/ijcai.2023/643>.

Coussemaker, Edmond de. *Scriptorum de Musica Medii Aevi*. Paris: A. Durand, 1864–1876.

Crawford, Tim and Richard Lewis. "Review: Music Encoding Initiative." *Journal of the American Musicological Society* 69 (1) (2016) : 273–285 <https://doi.org/10.1525/jams.2016.69.1.273>.

Creative Commons. *Attribution–ShareAlike 4.0 International — CC BY-SA 4.0*, 2013 <https://creativecommons.org/licenses/by-sa/4.0/>.

Crocker, Richard L. *A History of Musical Style*. New York, NY: McGraw-Hill, 1966.

Cross, Ian and Elizabeth Tolbert. "Epistemologies." *The Oxford Handbook of Western Music and Philosophy*. Eds. McAuley, Tomás, Nanette Nielsen, Jerrold Levinson, and Ariana Phillips-Hutton. Oxford: Oxford University Press, 2021. 265–282 <https://doi.org/10.1093/oxfordhb/9780199367313.013.14>.

Cuenca-Rodríguez, María Elena and Cory McKay. "Exploring musical style in the anonymous and doubtfully attributed mass movements of the Coimbra manuscripts: a statistical and machine learning approach.” *Journal of New Music Research* 50 (3) (2021) : 199–219 <https://doi.org/10.1080/09298215.2020.1870505>.

Cumming, Julie E., Cory McKay, Jonathan Stuchbery, and Ichiro Fujinaga. “Methodologies for Creating Symbolic Corpora of Western Music before 1600.” *Proceedings of the 19th International Conference on Music Information Retrieval (ISMIR)* (2018) : 491–498 <https://ismir2018.ismir.net/doc/pdfs/46_Paper.pdf>.

Cuthbert, Michael Scott and Christopher T. Ariza. “music21: A Toolkit for Computer-Aided Musicology and Symbolic Music Data.” *Proceedings of the 11th International Conference on Music Information Retrieval (ISMIR)* (2010) : 637–642 <http://ismir2010.ismir.net/proceedings/ismir2010-108.pdf>.

Damerau, Frederick J. “A technique for computer detection and correction of spelling errors.” *Communications of the ACM* 7 (3) (1964) : 171–176 <https://doi.org/10.1145%2F363958.363994>.

Dammann, Rolf. “Geschichte der Begriffsbestimmung Motette.” *Archiv für Musikwissenschaft* 16 (4) (1959) : 337–377 <https://doi.org/10.2307/930075>.

Dannenberg, Roger B. “A Structure for Representing, Displaying and Editing Music.” *Proceedings of the 1986 International Computer Music Conference* (1986) : 153–160 <https://www.cs.cmu.edu/˜rbd/subjbib2.html#Structure-ICMC-1986>.

DB-Engines. *DB-Engines Ranking*, 2025 <https://archive.ph/Cao4Q>.

Desmond, Karen. *Music and the moderni, 1300–1350: The ars nova in Theory and Practice*. Cambridge: Cambridge University Press, 2018 <https://doi.org/10.1017/9781316711545>.

———. “Notations.” *A Critical Companion to Medieval Motets*. Ed. Hartt, Jared C. Woodbridge: The Boydell Press, 2018. 103–130 <https://doi.org/10.1017/9781787442542.004>.

Desmond, Karen, Martha E. Thomae, Elizabeth Koshelev, Laurent Pugin, Emily Hopkins, Sam Howes, Sadie Menicanin, Shawn Mikkelson, Daniel Shapiro, and Arya Boudaie. *Measuring Polyphony: Digital Encodings of Late Medieval Music*, 2019– <https://measuringpolyphony.org/>.

Diaz, Chris. “Using static site generators for scholarly publications and open educational resources.” *Code4Lib Journal* 42 (2018) <https://journal.code4lib.org/articles/13861>.

Dick, Jeremy, Elizabeth Hull, and Ken Jackson. *Requirements Engineering*. 4th ed. New York, NY: Springer, 2017 <https://doi.org/10.1007/978-3-319-61073-3>.

Dittmer, Luther A. *Anonymous IV: Concerning the Measurement of Polyphonic Song*. New York, NY: Institute of Mediaeval Music, 1959a.

———. *Eine zentrale Quelle der Notre-Dame Musik. A central source of Notre-Dame* *polyphony*. New York, NY: Institute of Mediaeval Music, 1959b.

———. *Faksimile-Ausgabe der Handschrift, Facsimile Reproduction of the Manuscript Firenze, Biblioteca Mediceo-Laurenziana, Pluteo 29, 1*. New York, NY: Institute of Mediaeval Music, 1967.

———. *Faksimile-Ausgabe der Handschrift Wolfenbüttel 1099 Helmstadiensis-(1206), W2*. New York, NY: Institute of Mediaeval Music, 1969.

Dumitrescu, Theodor. "Ancient Concerns for the Twenty-First Century: 'Music Philology' in the Realm of the Digital." *Early Music Editing: Principles, Historiography, Future Directions*. Eds. Dumitrescu, Theodor, Karl Kügle, and Marnix van Berchum. Turnhout: Brepols, 2013.

Earp, Lawrence. "Notation II." *The Cambridge History of Medieval Music*. Eds. Everist, Mark and Thomas Forrest Kelly. 2. Cambridge: Cambridge University Press, 2018. 674–717 <https://doi.org/10.1017/9780511979866.023>.

Edmunds, Ben. *Securing PHP Apps*. New York, NY: Springer Science+Business Media, 2016 <https://doi.org/10.1007/978-1-4842-2120-4>.

Erickson, Raymond F. *Rhythmic Problems and Melodic Structure in Organum Purum: A Computer-Assisted Study*. PhD diss. New Haven, CT: Yale University, 1970.

Everist, Mark. "From Paris to St Andrews: The Origins of W_1." *Journal of the American Musicological Society* 43 (1) (1990) : 1–42 <https://doi.org/10.2307/831405>.

———. *Le Magnus Liber Organi de Notre-Dame de Paris. Volume II: Les Organa à Deux Voix pour l'Office du Manuscrit de Florence, Biblioteca Medicea-Laurenziana, Plut. 29.1*. Monaco: Éditions de L'Oiseau-Lyre, 2003.

———. *Discovering Medieval Song*. Cambridge: Cambridge University Press, 2018 <https://doi.org/10.1017/9780511852138>.

Falck, Robert. "Harmony in the Two-part Notre Dame Conductus." *Beyond the Moon: Festschrift Luther Dittmer*. Eds. Gillingham, Brian and Paul Merkley. Ottawa: The Institute of Mediaeval Music, 1990. 75–99.

Fétis, François-Joseph. *Histoire Générale de la Musique*. Paris: Librairie de Firmin Didot Frères, Fils et Cie, 1869.

Fielding, Roy Thomas. *Architectural Styles and the Design of Network-based Software Architectures*. PhD diss. Irvine, CA: University of California, 2000.

Flamary, Rémi, Nicolas Courty, Alexandre Gramfort, Mokhtar Z. Alaya, Aurélie Boisbunon, Stanislas Chambon, Laetitia Chapel, Adrien Corenflos, Kilian Fatras, Nemo Fournier, Léo Gautheron, Nathalie T. H. Gayraud, Hicham Janati, Alain Rakotomamonjy, Ievgen Redko, Antoine Rolet, Antony Schutz, Vivien Seguy, Danica J. Sutherland, Romain Tavenard, Alexander Tong, and Titouan Vayer. "POT: Python Optimal Transport." *Journal of Machine Learning Research* 22 (2021) : 1–8 <https://jmlr.org/papers/v22/20-451.html>.

Flotzinger, Rudolf. *Der Discantussatz im Magnus Liber und Seiner Nachfolge*. Wien: Böhlau, 1969.

———. “Clausula.” *The New Grove Dictionary of Music and Musicians*. Ed. Sadie, Stanley. 1st ed.. 4 London: Macmillan, 1980. 456–457 <https://doi.org/10.1093/gmo/9781561592630.article.05897>.

———. *Von Leonin zu Perotin: Der musikalische Paradigmenwechsel in Paris um 1210*. Bern: Peter Lang, 2007.

Flotzinger, Rudolf, Ernest H. Sanders, and Peter M. Lefferts. “Discant.” *The New Grove Dictionary of Music and Musicians*. 2nd ed.. 7 London: Macmillan, 2001. 366–374 <https://doi.org/10.1093/gmo/9781561592630.article.07839>.

Flowers, Andrew T. *Rhythm in the Polyphonic Conductus: A Computational Model and Its Implications*. PhD diss. Rochester, NY: University of Rochester, 2013 <http://hdl.handle.net/1802/27828>.

Fornés, Alicia, Anjan Dutta, Albert Gordo, and Josep Lladós. “CVC-MUSCIMA: a ground truth of handwritten music score images for writer identification and staff removal.” *International Journal on Document Analysis and Recognition (IJDAR)* 15 (2012) : 243–251 <https://doi.org/10.1007/s10032-011-0168-2>.

Foucault, Michel. *The Archaelogy of Knowledge*. Trans. A. M. Sheridan Smith. Abingdon: Routledge, 1972 <https://doi.org/10.4324/9780203604168>.

Fujinaga, Ichiro. *Adaptive Optical Music Recognition*. PhD diss. Montréal: McGill University, 1996 <https://escholarship.mcgill.ca/concern/theses/7w62f989f>.

———. “Staff Detection and Removal.” *Visual Perception of Music Notation: On-Line and Off-Line Recognition*. Ed. George, Susan E. London: IRM Press, 2005. 1–39 <https://doi.org/10.4018/978-1-59140-298-5.ch001>.

Fujinaga, Ichiro, Bo Alphonce, Bruce Pennycook, and Kharim Hogan. “Optical Music Recognition: Progress Report.” *Proceedings of the 1991 International Computer Music Conference* (1991) : 66–73 <http://hdl.handle.net/2027/spo.bbp2372.1991.013>.

Fuller, Sarah. “Early Polyphony to circa 1200.” *The Cambridge Companion to Medieval Music*. Ed. Everist, Mark. Cambridge: Cambridge University Press, 2011. 46–66 <https://doi.org/10.1017/ccol9780521846196.005>.

Gaspar, Drazena and Ivica Coric. *Bridging Relational and NoSQL Databases*. Advances in Data Mining and Database Management. Hershey, PA: IGI Global, 2018 <https://doi.org/10.4018/978-1-5225-3385-6>.

Girshick, Ross, Jeff Donahue, Trevor Darrell, and Jitendra Malik. “Rich feature hierarchies for accurate object detection and semantic segmentation.” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2014) : 580–587 <https://doi.org/10.48550/arXiv.1311.2524>.

Goodman, Joshua T. “A bit of progress in language modelling.” *Computer Speech and Language* 15 (4) (2001) : 403–434 <https://doi.org/10.1006/csla.2001.0174>.

Gouy, Manolo, Eric Tannier, Nicolas Conte, and David P. Parsons. “Seaview Version 5: A Multiplatform Software for Multiple Sequence Alignment, Molecular Phylogenetic Analyses, and Tree Reconciliation.” *Methods in molecular biology* 2231 (2021) : 241–260 <https://doi.org/10.1007/978-1-0716-1036-7_15>.

Grout, Donald Jay. *A History of Western Music*. 1st ed. New York, NY: W. W. Norton & Company, 1960.

Grout, Donald Jay and Claude V. Palisca. *A History of Western Music*. 6th ed. New York, NY: W. W. Norton & Company, 2001.

Haas, Max. “Organum.” *Musik in Geschichte und Gegenwart: Sachteil*. Ed. Finscher, Ludwig. 2nd ed.. 7 Kassel: Bärenreiter-Verlag, 1997. 853–881.

Haerder, Theo and Andreas Reuter. “Principles of Transaction-Oriented Database Recovery.” *ACM Computing Surveys* 15 (4) (1983) : 287–317 <https://doi.org/10.1145/289.291>.

Hajič Jr, Jan, Matthias Dorfer, Gerhard Widmer, and Pavel Pecina. “Towards Full-Pipeline Handwritten OMR with Musical Symbol Detection by U-Nets.” *Proceedings of the 19th International Conference on Music Information Retrieval (ISMIR)* (2018) : 225–232 <https://archives.ismir.net/ismir2018/paper/000175.pdf>.

Hall, Patrick A. V. and Geoff R. Dowling. “Approximate String Matching.” *ACM Computing Surveys* 12 (4) (1980) : 381–402 <https://doi.org/10.1145/356827.356830>.

Handschin, Jacques. “Eine wenig beachtete Stilrichtung innerhalb der mittelalterlichen Mehrstimmigkeit.” *Schweizerisches Jahrbuch für Musikwissenschaft* 1 (1924) : 56–75.

———. “Zur Frage der melodischen Paraphrasierung im Mittelalter.” *Zeitschrift für Musikwissenschaft* 10 (9) (1928) : 513–559.

———. “Der Organum-Traktat von Montpellier.” *Studien zur Musikgeschichte. Festschrift für Guido Adler zum 75. Geburtstag*. Wien: Universal-Edition, 1930. 50–57.

———. “Aus der alten Musiktheorie.” *Acta Musicologica* 14 (1942) : 1–27 <https://doi.org/10.2307/932087>.

Hankinson, Andrew. *Optical Music Recognition Infrastructure for Large-scale Music Document Analysis*. PhD diss. Montréal: McGill University, 2014 <https://escholarship.mcgill.ca/concern/theses/8g84mq59r>.

Hankinson, Andrew, Gabriel Vigliensoni, John Ashley Burgyone, and Ichiro Fujinaga. “Creating a Large-Scale Searchable Digital Collection from Printed Music

Materials.” *Proceedings of the 21st International Conference on World Wide Web* (2012) : 903–908 <https://doi.org/10.1145/2187980.2188221>.

Harrison, Frank Ll. *Music in Medieval Britain*. London: Butler & Tanner Ltd, 1958.

Harrison, Warren. “Eating Your Own Dog Food.” *IEEE Software* 23 (3) (2006) : 5–7 <https://doi.org/10.1109/ms.2006.72>.

Haus, Goffredo and Maurizio Longari. “A Multi-Layered, Time-Based Music Description Approach Based on XML.” *Computer Music Journal* 29 (1) (2005) : 70–85 <https://doi.org/10.1162/comj.2005.29.1.70>.

Hellsten, Alex. *Proposal: Use of D-clef in violin music*, 2016 <https://hellsten.kapsi.fi/dclef/>.

Hewlett, Walter B. and Eleanor Selfridge-Field. *Directory of Computer Assisted Research in Musicology 1987*, 1987 <https://www.ccarh.org/publications/books/cm/vol/3/contents.html>.

Hiley, David and Thomas B. Payne. “Notation §III, 2. Polyphony and secular monophony to c.1260.” *The New Grove Dictionary of Music and Musicians*. 2nd ed.. 18 London: Macmillan, 2001. 119–129 <https://doi.org/10.1093/gmo/9781561592630.article.20114>.

Hofmann, Klaus. “Zur Entstehungs- und Frühgeschichte des Terminus Motette.” *Acta Musicologica* 42 (3/4) (1970) : 138–150 <https://doi.org/10.2307/932192>.

Hoppin, Richard. *Medieval Music*. London: W. W. Norton & Company, 1978.

Huang, Xuedong, Fileno Alleva, Hsiao-Wuen Hon, Mei-Yuh Hwang, Kai-Fu Lee, and Ronald Rosenfeld. “The SPHINX-II speech recognition system: an overview.” *Computer Speech and Language* 7 (2) (1993) : 137–148 <https://doi.org/10.1006/csla.1993.1007>.

Hughes, Anselm. “Music in Fixed Rhythm.” *Early Medieval Music up to 1300*. Ed. Hughes, Anselm. New Oxford History of Music. 1st ed.. 2 London: Oxford University Press, 1954. 311–352.

Huron, David. “Design Principles in Computer-Based Music Representation.” *Computer Representations and Models in Music*. Eds. Marsden, Alan and Anthony Pople. Cambridge, MA: Academic Press, 1992. 5–39.

———. *The Humdrum Toolkit Reference Manual*. Menlo Park, CA: Center for Computer Assisted Research in the Humanities, 1995 <https://ccarh.org/publications/manuals/humdrumreference/humdrumreference-1995.pdf>.

———. *Music Research Using Humdrum: A User’s Guide*. Menlo Park, CA: Center for Computer Assisted Research in the Humanities, 1998 <http://ccarh.org/publications/manuals/humdrumuserguide/humdrum-users-guide-1998.pdf>.

Husmann, Heinrich and Gilbert Reaney. “The Origin and Destination of the ‘Magnus liber organi’.” *The Musical Quarterly* 49 (3) (1963) : 311–330

<https://doi.org/10.1093/mq/xlix.3.311>.

IEEE. "IEEE Recommended Practice for Speech Quality Measurements." "IEEE Recommended Practice for Speech Quality Measurements." *IEEE Transactions on Audio and Electroacoustics* 17 (3) (1969) : 225–246 <https://doi.org/10.1109%2FTAU.1969.1162058>.

Immel, Stephen C. "The Vatican Organum Treatise Re-Examined." *Early Music History* 20 (2001) : 121–172 <https://doi.org/10.1017/s0261127901001036>.

Intellectual Property Office. *Copyright notice: digital images, photographs and the internet*, 2021 <https://www.gov.uk/government/publications/copyright-notice-digital-images-photographs-and-the-internet/copyright-notice-digital-images-photographs-and-the-internet>.

International Music Score Library Project. *Magnus Liber Organi*, 2013 <https://imslp.org/wiki/Magnus_Liber_Organi_(Various)>.

ISO. *ISO 8879:1986: Information processing — Text and office systems — Standard Generalized Markup Language*, 1986 <http://www.iso.org/iso/catalogue_detail.htm?csnumber=16387>.

Jensenius, Alexander Refsum. "Best versus Good Enough Practices for Open Music Research." *Empirical Musicology Review* 16 (1) (2021) <https://doi.org/10.18061/emr.v16i1.7646>.

Jeong, Dasaem, Taegyun Kwon, Yoojin Kim, and Juhan Nam. "Graph Neural Network for Music Score Data and Modeling Expressive Piano Performance." *Proceedings of the 36th International Conference on Machine Learning* 97 (2019) : 3060–3070 <https://proceedings.mlr.press/v97/jeong19a.html>.

Jiang, Dejun, Zhenxing Wu, Chang-Yu Hsieh, Guangyong Chen, Ben Liao, Zhe Wang, Chao Shen, Dongsheng Cao, Jian Wu, and Tingjun Hou. "Could graph neural networks learn better molecular representation for drug discovery? A comparison study of descriptor-based and graph-based models." *Journal of Cheminformatics* 13 (12) (2021) <https://doi.org/10.1186/s13321-020-00479-8>.

Joiner, Richard E. "Gregorian Chant." *Computer studies in the humanities and verbal behavior* 2 (4) (1969) : 213–219.

Kania, Andrew. "Platonism vs. Nominalism in Contemporary Musical Ontology." *Art & Abstract Objects*. Ed. Mag Uidhir, Christy. Oxford: Oxford University Press, 2012 <https://doi.org/10.1093/acprof:oso/9780199691494.003.0010>.

———. "The Philosophy of Music." *The Stanford Encyclopedia of Philosophy* (2023) <https://plato.stanford.edu/archives/spr2023/entries/music/>.

Karp, Theodore. "Towards a Critical Edition of Notre Dame Organa Dupla." *The Musical Quarterly* 52 (3) (1966) : 350–367 <https://doi.org/10.1093/mq/lii.3.350>.

———. *The Polyphony of Saint Martial and Santiago de Compostela*. London: Oxford University Press, 1992.

Karystinaios, Emmanouil and Gerhard Widmer. "GraphMuse: A Library for Symbolic Music Graph Processing." *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)* (2024) <https://doi.org/10.48550/arXiv.2407.12671>.

Kassler, Michael. "A Sketch of the Use of Formalized Languages for the Assertion of Music." *Perspectives of New Music* 1 (2) (1963) : 83–94 <https://doi.org/10.2307/832106>.

Katoh, Kazutaka and Daron M. Standley. "MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability." *Molecular Biology and Evolution* 30 (4) (2013) : 772–780 <https://doi.org/10.1093/molbev/mst010>.

Kerman, Joseph. *Contemplating Music: Challenges to Musicology*. Cambridge, MA: Harvard University Press, 1985.

Kidwell, Susan A. "The Selection of Clausula Sources for Thirteenth-Century Motets: Some Practical Considerations and Aesthetic Implications)." *Current Musicology* 64 (1998) : 73–103 <https://journals.library.columbia.edu/index.php/currentmusicology/issue/view/511/CurrentMusicology64>.

Kierkegaard, Søren. *The Essential Kierkegaard*. Eds. Hong, Howard V. and Edna H. Hong. Princeton, NJ: Princeton University Press, 1978.

Kipf, Thomas N. and Max Welling. "Semi-supervised Classification with Graph Convolutional Networks." *IEEE Transactions on Neural Networks* 5 (1) (2016) : 61–80 <https://doi.org/10.1109/TNN.2008.2005605>.

Kullback, Solomon. "Information, Kullback." *Encyclopedia of Statistical Sciences*. Ed. Kotz, Samuel. 2nd ed.. 5 Hoboken, NJ: Wiley, 2006. 3505–3509 <https://doi.org/10.1002/0471667196.ess1372.pub2>.

Kurkela, Vesa and Markus Mantere. "Introduction." *Critical Music Historiography: Probing Canons, Ideologies and Institutions*. Eds. Kurkela, Vesa and Markus Mantere. Ashgate: Farnham, 2015. 1–13 <https://doi.org/10.4324/9781315575049>.

Leavitt, Neal. "Will NoSQL Databases Live Up to Their Promise?" *Computer* 43 (2) (2010) : 12–14 <https://doi.org/10.1109/MC.2010.58>.

Leech-Wilkinson, Daniel. *The Modern Invention of Medieval Music: Scholarship, Ideology, Performance*. Cambridge: Cambridge University Press, 2007.

Leichtentritt, Hugo. *Geschichte der Motette*. Leipzig: Breitkopf & Härtel, 1908.

Leman, Marc. "Symbolic and Subsymbolic Description of Music." *Music Processing*. Ed. Haus, Goffredo. Oxford: Oxford University Press, 1993. 119–164.

Levenshtein, Vladimir I. “Binary Codes Capable of Correcting Deletions, Insertions, and Reversals.” *Doklady Akademii Nauk SSSR* 163 (4) (1965) : 845–848.

Lexicon musicum Latinum medii aevi. *Clausula*, n.d. <https://woerterbuchnetz.de/?sigle=LmL&lemid=C00373>.

Li, Xiao, Li Sun, Mengjie Ling, and Yan Peng. “A survey of graph neural network based recommendation in social networks.” *Neurocomputing* 549 (126441) (2023) <https://doi.org/10.1016/j.neucom.2023.126441>.

Lim, Wen Qing, Jinhua Liang, and Huan Zhang. “Hierarchical Symbolic Pop Music Generation with Graph Neural Networks.” *eprint arXiv:2409.08155* (2024) <https://doi.org/10.48550/arXiv.2409.08155>.

Lindsay, Adam and Werner Kriechbaum. “There’s More Than One Way To Hear It: Multiple Representations of Music in MPEG-7.” *Journal of New Music Research* 4 (1999) : 364–372.

Liu, Xiaoran, Hang Yan, Shuo Zhang, Chenxin An, Xipeng Qiu, and Dahua Lin. “Scaling Laws of RoPE-based Extrapolation.” *Proceedings of the 12th International Conference on Learning Representations (ICLR)* (2024) <https://arxiv.org/abs/2310.05209>.

Lojkine, Ophir. *Dezoomify*, n.d. <https://dezoomify.ophir.dev/>.

Lord, Albert B. *The Singer of Tales*. Cambridge, MA: Harvard University Press, 1960.

Losseff, Nicky. *The Best Concords: Polyphonic Music in Thirteenth-Century Britain*. Outstanding Dissertations in Music from British Universities. London: Garland Publishing, 1994.

Lu, Jie, Anjin Liu, Fan Dong, Feng Gu, João Gama, and Guangquan Zhang. “Learning under Concept Drift: A Review.” *IEEE Transactions on Knowledge and Data Engineering* 31 (12) (2019) <https://doi.org/10.1109/TKDE.2018.2876857>.

Ludwig, Friedrich. “Studien über die Geschichte der mehrstimmigen Musik im Mittelalter. I. Die mehrstimmige Musik der ältesten Epoch im Dienste der Liturgie.” *Kirchenmusikalisches Jahrbuch* 19 (1905) : 1–10.

———. “Studien über die Geschichte der mehrstimmigen Musik im Mittelalter. III. Über die Entstehung und die erste Entwicklung der lateinischen und französischen Motette in musikalischer Beziehung.” *Sammelbände der International Musikgesellschaft* 7 (1906) : 514–528.

———. “Die liturgischen Organa Leonins und Perotins.” *Riemann Festschrift: Gesammelte Studien. Hugo Riemann zum sechzigsten Geburtstag überreicht von Freunden und Schülern*. Leipzig: Max Hesse-Verlag, 1909. 200–213.

———. *Repertorium Organum Recentioris et Motetorum Vetustissimi Stili*. Ed. Dittmer, Luther A. New York, NY: Institute of Mediaeval Music, 1910.

Makarov, Ilya, Dmitrii Kiselev, Nikita Nikitinsky, and Lovro Subelj. “Survey on graph embeddings and their applications to machine learning problems on graphs." *PeerJ Computer Science* 7 (2021) <https://doi.org/10.7717/peerj-cs.357>.

Marino, Mark C. *Critical Code Studies*. Cambridge, MA: The MIT Press, 2020 <https://doi.org/10.7551/mitpress/12122.001.0001>.

Martin, Nathan John. "Corpus Studies and 'Close Listening'." *Music Analysis* 43 (2) (2024) : 191–246 <https://doi.org/10.1111/musa.12229>.

Matheson, Carl and Ben Caplan. "Ontology." *The Routledge Companion to Philosophy and Music*. Eds. Gracyk, Theodore and Andrew Kania. Abingdon: Routledge, 2011. 37–47.

Mathias, Adam John. *Creating Clausulae at Notre-Dame-de-Paris: A Study of Compositional Processes and Techniques*. PhD diss. Cambridge: University of Cambridge, 2020 <https://repository.cam.ac.uk/handle/1810/311353>.

Maxwell, John Turner. *Mockingbird: An Interactive Composer's Aid*. Masters diss. Cambridge, MA: Massachusetts Institute of Technology, 1981.

McGarry, Daniel D. *The Metalogicon of John of Salisbury: A Twelfth-Century Defense of the Verbal and Logical Arts of the Trivium*. Gloucester: MA: Peter Smith, 1971.

McKinley, Dan. *Choose Boring Technology*, 2015 <https://mcfunley.com/choose-boring-technology>.

McLuhan, Marshall. *Understanding Media: the Extensions of Man*. New York: McGraw-Hill Book Company, 1964.

Meyer, Wilhelm. "Der Ursprung des Motetts." *Gesammelte Abhandlungen zur Mittellateinischen Rhythmik*. 1898. 2. Berlin: Weidmannsche Buchhandlung, 1905. 303–341.

Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. "Efficient Estimation of Word Representations in Vector Space." *eprint arXiv:1301.3781* (2013) <https://doi.org/10.48550/arXiv.1301.3781>.

Morent, Stefan. "Representing a Medieval Repertory and its Sources: The Music of Hildegard von Bingen." *Computing in Musicology* 12 (2000) : 19–33.

Moretti, Franco. *Distant Reading*. London: Verso, 2013.

Morgan, Alexander, Daniel Russo-Batterham, and Richard Freedman. "Musicologists and Data Scientists Pull out all the Stops: Defining Renaissance Cadences Systematically." [Paper presentation]. *Music Encoding Conference, 19–22 May 2022, Halifax, Canada* (2022) <https://sites.google.com/haverford.edu/crim-project/events/music-encoding-conference-2022>.

Müllensiefen, Daniel and Klaus Frieler. "Statistical Methods in Music Corpus Studies: Application, Use Cases, and Best Practice Examples." *The Oxford Handbook of Music and Corpus Studies*. Eds. Shanahan, Daniel, John Ashley Burgoyne, and Ian Quinn. Oxford: Oxford University Press, 2022

<https://doi.org/10.1093/oxfordhb/9780190945442.001.0001>.

Needleman, Saul B. and Christian D. Wunsch. “A general method applicable to the search for similarities in the amino acid sequence of two proteins.” *Journal of Molecular Biology* 48 (3) (1970) : 443–453 <https://doi.org/10.1016%2F0022-2836%2870%2990057-4>.

Nemarich, Eric. “Organistae and the Cultivation of Polyphony at Notre-Dame de Paris, c. 1190–1273.” *Speculum* 99 (1) (2024) : 102–142 <https://doi.org/10.1086/727887>.

Nettl, Bruno. *The Study of Ethnomusicology: Thirty-one Issues and Concepts*. 2nd ed. Champaign, IL: University of Illinois Press, 2005.

Neuwirth, Markus and Martin Rohrmeier. “Wie wissenschaftlich muss Musiktheorie sein? Chancen und Herausforderungen musikalischer Korpusforschung.” *Zeitschrift der Gesellschaft für Musiktheorie* 13 (2) (2016) : 171–193 <https://doi.org/10.31751/915>.

Nooshin, Laudan. “Introduction to the Special Issue: The Ethnomusicology of Western Art Music.” *Ethnomusicology Forum* 20 (3) (2011) : 285–300 <https://doi.org/10.1080/17411912.2011.659439>.

Notredame, Cédric, Desmond G Higgins, and Jaap Heringa. “T-coffee: a novel method for fast and accurate multiple sequence alignment.” *Journal of Molecular Biology* 302 (1) (2000) : 205–217 <https://doi.org/10.1006/jmbi.2000.4042>.

Nova, Oz. *You are Not Google*, 2017 <https://blog.bradfieldcs.com/you-are-not-google-84912cf44afb>.

Nuttall, Thomas, Miguel García Casado, Victor Núñez Tarifa, Rafael Caro Repetto, and Xavier Serra. “Contributing to new musicological theories with computational methods: The case of centonization in Arab-Andalusian music.” *Proceedings of the 20th International Conference on Music Information Retrieval (ISMIR)* (2019) : 223–228 <https://archives.ismir.net/ismir2019/paper/000025.pdf>.

Owens, Michael. *The Definitive Guide to SQLite*. New York, NY: Springer-Verlag, 2006.

Pace, Ian. “Notation, Time and the Performer’s Relationship to the Score in Contemporary Music.” *Unfolding Time: Studies in Temporality in Twentieth-Century Music*. Ed. Crispin, Darla. Leuven: Leuven University Press, 2009. 149–191 <https://doi.org/10.2307/j.ctt9qdxmr.8>.

Pacherie, Elisabeth. “Do We See with Microscopes?” *The Monist* 78 (2) (1995) : 171–188 <https://doi.org/10.5840/monist199578217>.

Page, Christopher. *The Owl and the Nightingale: Musical Life and Ideas in France 1100–1300*. London: J.M. Dent & Sons, 1989.

———. *Discarding Images: Reflections on Music & Culture in Medieval France*.

Oxford: Oxford University Press, 1993.

Page, Roderic. "Liberating links between datasets using lightweight data publishing: an example using plant names and the taxonomic literature." *Biodiversity Data Journal* 6 (2018) <https://doi.org/10.3897/BDJ.6.e27539>.

Pang, Yanwei and Jiale Cao. "Deep Learning in Object Detection." *Deep Learning in Object Detection and Recognition*. Eds. Jiang, Xiaoyue, Abdenour Hadid, Yanwei Pang, Eric Granger, and Xiaoyi Feng. Singapore: Springer Nature, 2019. 19–57 <https://doi.org/10.1007/978-981-10-5152-4_2>.

Parrish, Carl. *The Notation of Medieval Music*. New York, NY: W. W. Norton & Company, 1957.

Parry, C. Hubert H. *The Art of Music*. London: Kegan Paul, Trench, Trübner, & Co. Ltd, 1893.

Paul, Ashis, Rishav Pramanik, Samir Malakar, and Ram Sarkar. "An ensemble of deep transfer learning models for handwritten music symbol recognition." *Neural Computing and Applications* 34 (2022) : 10409–10427 <https://doi.org/10.1007/s00521-021-06629-9>.

Payne, Thomas B. *Le Magnus Liber Organi de Notre-Dame de Paris. Volume VI: Les Organa à Deux Voix du Manuscrit de Wolfenbüttel, Hertzog August Bibliothek, Cod. Guelf. 1099 Helmst.* Monaco: Éditions de L'Oiseau-Lyre, 1996.

———. *The Organa, Clausula and Conductus from MS I-Fl Pluteus 29.1*. Digital Image Archive of Medieval Music, 2021– <https://www.diamm.ac.uk/resources/music-editions/organa-and-clausulae-ms-i-fl-pluteus-291/>.

Pesce, Dolores. "Theory and notation." *The Cambridge Companion to Medieval Music*. Ed. Everist, Mark. Cambridge: Cambridge University Press, 2011. 276–290 <https://doi.org/10.1017/ccol9780521846196.018>.

Peter, Silvan David and Gerhard Widmer. "TheGlueNote: learned representations for robust and flexible note alignments." *Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR)* (2024) <https://doi.org/10.48550/arXiv.2408.04309>.

Planchart, Alejandro Enrique. "The Flower's Children." *Journal of Musicological Research* 22 (4) (2003) : 303–348 <https://doi.org/10.1080/714040828>.

Presner, Todd. *Digital Humanities 2.0: A Report on Knowledge*, 2010 <https://web.archive.org/web/20120418110756/http://cnx.org/content/m34246/1.6/?format=pdf>.

Pugin, Laurent. "Optical Music Recognition of Early Typographic Prints using Hidden Markov Models." *Proceedings of the 7th International Conference on Music Information Retrieval (ISMIR)* (2006) <https://ismir2006.ismir.net/PAPERS/ISMIR06152_Paper.pdf>.

Pugin, Laurent, Jason Hockman, John Ashley Burgoyne, and Ichiro Fujinaga. “Gamera versus Aruspix: Two Optical Music Recognition Approaches.” *Proceedings of the 9th International Conference on Music Information Retrieval (ISMIR)* (2008) <https://ismir2008.ismir.net/papers/ISMIR2008_247.pdf>.

Quinn, Ian. “Review of ‘Hidden Structure: Music Analysis Using Computers’ and ‘Music21: A Toolkit for Computer-Aided Musicology’.” *Journal of the American Musicological Society* 67 (1) (2014) : 295–307 <https://doi.org/10.1525/jams.2014.67.1.295>.

Ramsay, Stephen. *Reading Machines: Toward an Algorithmic Criticism*. Urbana, IL: University of Illinois Press, 2011 <https://doi.org/10.5406/illinois/9780252036415.001.0001>.

———. “On Building.” *Defining Digital Humanities: A Reader*. Eds. Terras, Melissa, Julianne Nyhan, and Edward Vanhoutte. Farnham: Ashgate, 2013. 243–245.

Rankin, Susan. “Review: Parisian Polyphony.” *Early Music* 25 (1) (1997) : 139–141 <https://doi.org/10.1093/earlyj/xxv.1.139>.

Rebelo, Ana, Ichiro Fujinaga, Filipe Paszkiewicz, Andre R. S. Marcal, Carlos Guedes, and Jaime S. Cardoso. “Optical music recognition: state-of-the-art and open issues.” *International Journal of Multimedia Information Retrieval* 1 (2012) : 173–190 <https://doi.org/10.1007/s13735-012-0004-6>.

Rebelo, Ana, Jakub Tkaczuk, Ricardo Sousa, and Jaime S. Cardoso. “Metric Learning for Music Symbol Recognition.” *Proceedings of the 10th International Conference on Machine Learning and Applications* 2 (2011) : 106–111 <https://doi.org/10.1109/ICMLA.2011.94>.

Reckow, Fritz. *Der Musiktraktat des Anonymus 4. Teil II: Interpretation der Organum Purum-lehre*. Wiesbaden: Franz Steiner Verlag, 1967.

———. “Das Organum.” *Gattungen der Musik in Einzeldarstellungen: Gedenkschrift Leo Schrade*. Ed. Arlt, Wulf. Bern: Francke, 1973. 434–496.

Redmon, Joseph, Santosh Divvala, Ross Girshick, and Ali Farhadi. “You Only Look Once: Unified, Real-Time Object Detection.” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016) : 779–788 <https://doi.org/10.48550/arXiv.1506.02640>.

Reese, Gustave. *Music in the Middle Ages*. New York, NY: W. W. Norton & Company, 1940.

Ricci, Massimo Masani. *Codice Pluteo 29.1 della Biblioteca Laurenziana di Firenze: Storia e catalogo comparato*. Pisa: Edizioni ETS, 2002.

Ríos-Vila, Antonio, Jorge Calvo-Zaragoza, David Rizo, and José M. Iñesta. “ReadSco: An Open-Source Web-Based Optical Music Recognition Tool.” *Proceedings of the 2nd International Workshop on Reading Music Systems* (2019) : 27–30

<https://arxiv.org/pdf/2212.00380.pdf#page=27>.

Rizo, David, Jorge Calvo-Zaragoza, and José M. Iñesta. “MuRET: A music recognition, encoding, and transcription tool.” *Proceedings of the 5th International Conference on Digital Libraries for Musicology (DLfM)* (2018) : 52–56 <https://doi.org/10.1145/3273024.3273029>.

Rodin, Jesse and Craig Sapp. *The Josquin Research Project*, 2010– <https://josquin.stanford.edu/>.

Roesner, Edward H. *The Manuscript Wolfenbüttel, Herzog-August-Bibliothek, 628 Helmstadiensis: A Study of its Origins and of its Eleventh Fascicle*. PhD diss. New York, NY: New York University, 1974.

———. “The Origins of W_1.” *Journal of the American Musicological Society* 29 (3) (1976) : 337–380 <https://doi.org/10.2307/830966>.

———. “The Performance of Parisian Organum.” *Early Music* 7 (2) (1979) : 174–189 <https://doi.org/10.1093/earlyj/7.2.174>.

———. “The problem of chronology in the transmission of organum duplum.” *Music in Medieval and Early Modern Europe: Patronage, Sources and Texts*. Ed. Fenlon, Iain. Cambridge: Cambridge University Press, 1981. 365–399.

———. *Le Magnus Liber Organi de Notre-Dame de Paris. Volume I: Les quadrupla et tripla de Paris*. Monaco: Éditions de L’Oiseau-Lyre, 1993.

———. “Who ‘Made’ the Magnus Liber?” *Early Music History* 20 (2001) : 227–266 <https://doi.org/10.1017/S0261127901001061>.

———. *Le Magnus Liber Organi de Notre-Dame de Paris. Volume VII: Les Organa et les Clausules à Deux Voix du Manuscrit de Wolfenbüttel, Herzog August Bibliothek, Cod. Guelf. 628 Helmst.* Monaco: Éditions de L’Oiseau-Lyre, 2009.

———. “Notre Dame.” *The Cambridge History of Medieval Music*. Eds. Everist, Mark and Thomas Forrest Kelly. 2. Cambridge: Cambridge University Press, 2018. 834–880 <https://doi.org/10.1017/9780511979866.028>.

Roland, Perry, Andrew Hankinson, and Laurent Pugin. “Early music and the Music Encoding Initiative.” *Early Music* 42 (4) (2014) : 605–611 <https://doi.org/10.1093/em/cau098>.

Roth-Burnette, Jennifer Louise. *Organizing Scripture: Organum Melos, Composition, and Memoria in a Group of Notre-Dame Responsories*. PhD diss. New York, NY: New York University, 2010.

Saleh, Zeyad, Ké Zhang, Jorge Calvo-Zaragoza, Gabriel Vigliensoni, and Ichiro Fujinaga. “Pixel.js: Web-based Pixel Classification Correction Platform for Ground Truth Creation.” *Proceedings of the 14th IAPR International Conference on Document Analysis and Recognition* (2017) : 39–40 <https://doi.org/10.1109/ICDAR.2017.267>.

Sanders, Ernest H. “Peripheral Polyphony of the 13th Century.” *Journal of the American Musicological Society* 17 (3) (1964) : 261–287 <https://doi.org/10.2307/830091>.

———. “The Question of Perotin’s Oeuvre and Dates.” *Festschrift für Walter Wiora: zum 30.Dezember 1966*. Eds. Finscher, Ludwig and Christoph-Hellmut Mahling. Kassel: Bärenreiter, 1967. 241–249.

———. “The Medieval Motet.” *Gattungen der Musik in Einzeldarstellungen: Gedenkschrift Leo Schrade*. Eds. Arlt, Wulf, Ernst Lichtenhahn, Hans Oesch, and Max Haas. Bern: Francke, 1973. 497–573.

———. “Consonance and Rhythm in the Organum of the 12th and 13th Centuries.” *Journal of the American Musicological Society* 33 (2) (1980) : 264–286 <https://doi.org/10.2307/831113>.

———. “Motet, §I. Medieval.” *The New Grove Dictionary of Music and Musicians*. Ed. Sadie, Stanley. 1st ed.. 12 London: Macmillan, 1980. 617–628.

Schmalzriedt, Siegfried. “Clausula.” *Handwörterbuch Der Musikalischen Terminologie*. Ed. Eggebrecht, Hans Heinrich. 1. Stuttgart: Franz Steiner, 1974.

Schmidt, Helmut. “Zur Melodiebildung Leonins und Perotins.” *Zeitschrift für Musikwissenschaft* 14 (3) (1931) : 129–134.

Schram, Aaron and Kenneth M. Anderson. “MySQL to NoSQL: data modeling challenges in supporting scalability.” *SPLASH ’12: Proceedings of the 3rd annual conference on Systems, programming, and applications: software for humanity* (2012) : 191–202 <https://doi.org/10.1145/2384716.2384773>.

Schreiber, Jacob. “Pomegranate: Fast and Flexible Probabilistic Modeling in Python.” *Journal of Machine Learning Research* 18 (164) (2018) : 1–6 <https://jmlr.org/papers/v18/17-636.html>.

Schreiner, Geomar A., Denio Duarte, and Ronaldo dos Santos Mello. “When Relational-Based Applications Go to NoSQL Databases: A Survey.” *Information* 10 (7) (2019) <https://doi.org/10.3390/info10070241>.

Sears, David R. W. and Gerhard Widmer. “Beneath (or beyond) the surface: Discovering voice-leading patterns with skip-grams.” *Journal of Mathematics and Music* 15 (3) (2020) : 209–234 <https://doi.org/10.1080/17459737.2020.1785568>.

Selfridge-Field, Eleanor. *Beyond MIDI: The Handbook of Musical Codes*. Cambridge, MA: The MIT Press, 1997.

———. “Conceptual and Representational Issues in Melodic Comparison.” *Melodic Similarity: Concepts, Procedures, and Applications*. Eds. Hewlett, Walter B. and Eleanor Selfridge-Field. Cambridge, MA: The MIT Press, 1998. 3–64.

Shanahan, Daniel. “Empirical Musicology: An Interview with David Huron, Part I.”

Empirical Musicology Review 15 (1) (2020) <https://doi.org/10.18061/emr.v15i1-2.7718>.

Shannon, Claude E. “A Mathematical Theory of Communication.” *Claude E. Shannon: Collected Papers*. Eds. Sloane, N. J. A and Aaron D. Wyner. IEEE. 5–83 <https://doi.org/10.1109/9780470544242.ch1>.

Sievers, Fabian and Desmond G. Higgins. “Clustal Omega for making accurate alignments of many protein sequences.” *Protein Science* 27 (1) (2017) : 135–145 <https://doi.org/10.1002/pro.3290>.

Simon, Herbert A. *The Sciences of the Artificial*. 3rd ed. Cambridge, MA: The MIT Press, 1996 <https://doi.org/10.7551/mitpress/12107.001.0001>.

Simonyan, Karen and Andrew Zisserman. “Very Deep Convolutional Networks for Large-Scale Image Recognition.” [Paper presentation]. *International Conference on Learning Representations (ICLR)* (2015) <https://doi.org/10.48550/arXiv.1409.1556>.

Simsion, Graeme C. and Graham C. Witt. *Data Modeling Essentials*. 3rd ed. Oxford: Elsevier, 2005.

Sloan, Donald. “Aspects of Music Representation in HyTime/SMDL.” *Computer Music Journal* 17 (4) (1993) : 51–59 <https://doi.org/10.2307/3680544>.

Smiraglia, Richard P. *Domain Analysis for Knowledge Organization: Tools for Ontology Extraction*. Kidlington: Chandos, 2015 <https://doi.org/10.1016/c2014-0-02303-9>.

Smith, Norman E. *The Clausulae of the Notre Dame School: A Repertorial Study*. PhD diss. New Haven, CT: Yale University, 1964.

———. “Interrelationships among the Alleluias of the ‘Magnus liber organi’.” *Journal of the American Musicological Society* 25 (2) (1972) : 175–202 <https://doi.org/10.2307/830595>.

———. “Interrelationships among the Graduals of the Magnus Liber Organi.” *Acta Musicologica* 45 (1) (1973) : 73–97 <https://doi.org/10.2307/932223>.

———. “The Earliest Motets: Music and Words.” *Journal of the Royal Musical Association* 114 (2) (1989) : 141–163 <https://doi.org/10.1093/jrma/114.2.141>.

Smithies, James. “The Challenge of the Digital Humanities.” *The Digital Humanities and the Digital Modern*. London: Macmillan, 2015. 1–15 <https://doi.org/10.1057/978-1-137-49944-8_1>.

Sporny, Manu, Shane McCarron, Ben Adida, Mark Birbeck, Gregg Kellogg, Ivan Herman, and Steven Pemberton. *W3C Recommendation: HTML+RDFa 1.1*. 2nd ed., 2015 <https://www.w3.org/TR/html-rdfa/>.

Steck, Harald, Chaitanya Ekanadham, and Nathan Kallus. “Is Cosine-Similarity of Embeddings Really About Similarity?” *WWW ’24: Companion Proceedings of the*

ACM Web Conference 2024 (2024) : 887–890 <https://doi.org/10.1145/3589335.3651526>.

Steiner, Katherine Kennedy. "The Scribe of W_1 and His Scottish Context." *The Journal of Musicology* 38 (3) (2021) : 364–399 <https://doi.org/10.1525/JM.2021.38.3.364>.

Steyn, Jacques. *Structuring Music through Markup Language*. Hershey, PA: Information Science Reference, 2013.

Stokes, Simon. *Art and Copyright*. 3rd ed. Oxford: Bloomsbury, 2020 <https://doi.org/10.5040/9781509934287>.

Stutter, Joshua. *The musical, notational, and codicological evidence of W1 for an oral transmission of Notre Dame polyphony to Scotland*. Masters diss. Glasgow: University of Glasgow, 2020 <http://theses.gla.ac.uk/id/eprint/79015>.

———. "Annotation of Medieval Music Facsimiles Using 'Good Enough' OMR." *Proceedings of the Music Encoding Conference* (2021) <https://hcommons.org/deposits/item/hc:46013>.

———. "Organal motets, organum prosulae, polyphonic tropes, motet-likes: the universalisation of the motet." [Paper presentation]. *Benedicamus Domino conference, 7–8 September 2023, Oslo* (2023).

———. "Data structures for music encoding: tables, trees, and graphs." *Journal of New Music Research* (2024) <https://doi.org/10.1080/09298215.2024.2412598>.

Su, Jianlin, Yu Lu, Shengfeng Pan, Ahmed Murtadha, Bo Wen, and Yunfeng Liu. "RoFormer: Enhanced Transformer with Rotary Position Embedding." *Neurocomputing* 568 (2024) <https://doi.org/10.1016/j.neucom.2023.127063>.

Sunardi, Andri and Suharjito. "MVC Architecture: A Comparative Study Between Laravel Framework and Slim Framework in Freelancer Project Monitoring System Web Based." *Procedia Computer Science* 157 (2019) : 134–141 <https://doi.org/10.1016/j.procs.2019.08.150>.

Taruskin, Richard. *Music from the Earliest Notations to the Sixteenth Century*. Oxford History of Western Music. Vol. 1. London: Oxford University Press, 2005 <https://www.oxfordwesternmusic.com/view/Volume1/actrade-9780195384819-miscMatter-021008.xml>.

Text Encoding Initiative. *Linking, Segmentation, and Alignment*, 2025 <https://tei-c.org/release/doc/tei-p5-doc/en/html/SA.html>.

Thomae, Martha E. *Guatemalan Cathedral Choirbooks: From Manuscript to Digital Images to Edited Symbolic Scores*. PhD diss. Montréal: McGill University, 2022 <https://escholarship.mcgill.ca/concern/theses/d217qw26r>.

Tischler, Hans. "A propos a Critical Edition of the Parisian Organa Dupla." *Acta Musicologica* 40 (1) (1968) : 28–43 <https://doi.org/10.2307/932281>.

———. “How Were Notre Dame Clausulae Performed?” *Music & Letters* 50 (2) (1969) : 273–277 <https://doi.org/10.1093/ml/l.2.273>.

———. “The Evolution of the ‘Magnus Liber Organi’.” *The Musical Quarterly* 70 (2) (1984) : 163–174 <https://www.jstor.org/stable/742208>.

———. *The Parisian Two-part Organa: the Complete Comparative Edition*. Stuyvesant, NY: Pendragon Press, 1988.

Tomlinson, Gary. *Music in Renaissance Magic: Toward a Historiography of Others*. London: University of Chicago Press, 1993.

Tosh, John. *The Pursuit of History: Aims, methods and new direction in the study of modern history*. 3rd ed. London: Pearson Education, 1999.

Tóth-Czifra, Erzsébet. “The Risk of Losing the ‘Thick Description’: Data Management Challenges Faces by the Arts and Humanities in the Evolving FAIR Data Ecosystem.” *Digital Technology and the Practices of Humanities Research*. Ed. Edmond, Jennifer. Cambridge: Open Book Publishers, 2020 <https://doi.org/10.11647/obp.0192.10>.

Touvron, Hugo, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothee Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. “LLaMA: Open and Efficient Foundation Language Models.” *eprint arXiv:2302.13971* (2023) <https://arxiv.org/abs/2302.13971>.

Treitler, Leo. “Homer and Gregory: The Transmission of Epic Poetry and Plainchant.” *With Voice and Pen: Coming to Know Medieval Song and How it was Made*. Oxford: Oxford University Press, 2003d. 131–185 <https://doi.org/10.1093/acprof:oso/9780199214761.003.0006>.

———. “Preface.” *With Voice and Pen: Coming to Know Medieval Song and How it was Made*. Oxford: Oxford University Press, 2003a. vii–xiii <https://doi.org/10.1093/acprof:oso/9780199214761.002.0005>.

———. “The Vatican Organum Treatise and the Organum of Notre Dame of Paris: Perspectives on the Development of a Literature Music Culture in Europe.” *With Voice and Pen: Coming to Know Medieval Song and How it was Made*. Oxford: Oxford University Press, 2003b. 68–83 <https://doi.org/10.1093/acprof:oso/9780199214761.003.0003>.

———. “Written Music and Oral Music: Improvisation in Medieval Performance.” *With Voice and Pen: Coming to Know Medieval Song and How it was Made*. Oxford: Oxford University Press, 2003c. 39–67 <https://doi.org/10.1093/acprof:oso/9780199214761.003.0002>.

University of Pennyslvania. “Deaths: Dr. Smith, Music.” *University of Pennsylvania Almanac* 60 (26) (2014) : 2 <https://www.upenn.edu/almanac>.

Urberg, Michelle. “Pasts and Futures of Digital Humanities in Musicology: Moving Towards a ‘Bigger Tent’.” *Music Reference Services Quarterly* 20 (3) (2017) : 134–150 <https://doi.org/10.1080/10588167.2017.1404301>.

Ursprung, Otto. *Die Katholische Kirchenmusik*. Potsdam: Akademische Verlagsgesellschaft Athenaion, 1931.

van der Werf, Hendrik. *Integrated Directory of Organa, Clausulae, and Motets of the Thirteenth Century*. Rochester, NY: Self-published, 1989.

———. “Review of The Parisian Two-Part Organa: The Complete Comparative Edition by Hans Tischler.” *Notes* 47 (3) (1991) : 950–953 <https://doi.org/10.2307/941937>.

———. “Anonymous IV as chronicler.” *Musicology Australia* 15 (1) (1992) : 3–13 <https://doi.org/10.1080/08145857.1992.10415192>.

van Gompel, Maarten and Antal van den Bosch. “Efficient n-gram, Skipgram and Flexgram Modelling with Colibri Core.” *Journal of Open Research Software* 4 (2016) <https://doi.org/10.5334/jors.105>.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, and Łukasz Kaiser. “Attention Is All You Need.” *Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS)* (2017) <https://papers.nips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html>.

Villani, Cédric. *Optimal Transport: Old and New*. Berlin: Springer, 2009 <https://doi.org/10.1007/978-3-540-71050-9>.

Waite, William G. *The Rhythm of Twelfth-Century Polyphony*. New Haven, CT: Yale University Press, 1954.

———. “The Abbreviation of the Magnus Liber.” *Journal of the American Musicological Society* 14 (2) (1961) : 147–158 <https://doi.org/10.2307/829753>.

Wegman, Rob C. “Historical Musicology: Is It Still Possible?” .*The Cultural Study of Music: A Critical Introduction*. Eds. Clayton, Martin, Trevor Herbert, and Richard Middleton. 2nd ed. Abingdon: Routledge, 2012. 40–48 <https://doi.org/10.4324/9780203821015-17>.

———. “The World according to Anonymous IV.” *Qui musicam in se habet: Studies in Honor of Alejandro Enrique Planchart*. Eds. Zayaruznaya, Anna, Bonnie J. Blackburn, and Stanley Boorman. Middleton, WI: American Institute of Musicology, 2015. 693–729.

Weisser, Martin. *Practical Corpus Linguistics: An Introduction to Corpus-Based Language Analysis*. Chichester: Wiley, 2016.

Wen, Cuihong and Longjiao Zhu. “A Sequence-to-Sequence Framework Based on Transformer With Masked Language Model for Optical Music Recognition.” *IEEE*

Access 10 (2022) : 118243–118252 <https://doi.org/10.1109/ACCESS.2022.3220878>.

Wick, Christoph, Alexander Hartelt, and Frank Puppe. "Staff, Symbol and Melody Detection of Medieval Manuscripts Written in Square Notation Using Deep Fully Convolutional Networks." *Applied Sciences* 9 (13) (2019) <https://doi.org/10.3390/app9132646>.

Wick, Christoph and Frank Puppe. "OMMR4All: a Semiautomatic Online Editor for Medieval Music Notations." *Proceedings of the 2nd International Workshop of Reading Music Systems* (2019) : 31–34 <https://arxiv.org/pdf/2212.00380.pdf#page=32>.

Wiggins, Geraint A. "Computer Representation of Music in the Research Environment." *Modern Methods for Musicology: Prospects, Proposals, and Realities*. Eds. Crawford, Tim and Lorna Gibson. Surrey: Ashgate, 2009a. 7–22.

———. "Semantic Gap?? Schemantic Schmap!! Methodological Considerations in the Scientific Study of Music." *Proceedings of the 11th IEEE International Symposium on Multimedia* (2009b) : 477–482 <https://doi.org/10.1109/ism.2009.36>.

Wilkinson, Mark D., Michel Dumontier, Ijsbrand Jan Aalsberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E. Bourne, Jildau Bouwman, Anthony J. Brookes, Tim Clark, Mercè Crosas, Ingrid Dillo, Olivier Dumon, Scott Edmunds, Chris T. Evelo, Richard Finkers, Alejandra Gonzalez-Beltran, …, and Barend Mons. "The FAIR Guiding Principles for scientific data management and stewardship." *Scientific Data* 3 (2016) <https://doi.org/10.1038/sdata.2016.18>.

Wolfenbütteler Digitale Bibliothek, . *Nutzungsbedingungen für Online-Angebote der Herzog August Bibliothek Wolfenbüttel*, 2013 <https://diglib.hab.de/copyright.html>.

Wolfenbütteler Digitale Bibliothek. *WDB: mss/1099-helmst*, n.d. <https://diglib.hab.de/wdb.php?dir=mss/1099-helmst>.

Wolfenbütteler Digitale Bibliothek. *WDB: mss/628-helmst*, n.d. <https://diglib.hab.de/wdb.php?dir=mss/628-helmst>.

Wooldridge, H. E. "Discant or Measured Music." *Oxford History of Music: The Polyphonic Period. Part I.* 1. London: University of Oxford, 1901. 102–388.

Wright, Craig. "Leoninus, Poet and Musician." *Journal of the American Musicological Society* 39 (1) (1986) : 1–35 <https://doi.org/10.2307/831693>.

———. *Music and ceremony at Notre Dame of Paris, 500–1500*. Cambridge: Cambridge University Press, 1989.

Yang, Junwen, Pranav Subramaniam, Shan Lu, Cong Yan, and Alvin Cheung. "How not to structure your database-backed web applications: a study of performance bugs in the wild." *Proceedings of the 40th International Conference on Software Engineering* (2018) : 800–810 <https://doi.org/10.1145/3180155.3180194>.

Yudkin, Jeremy. "The 'Copula' According to Johannes de Garlandia." *Musica Disciplina* 34 (1980) : 67–84 <https://doi.org/10.4324/9781315097015-7>.

———. "The Rhythm of Organum Purum." *The Journal of Musicology* 2 (1983) : 355–376 <https://doi.org/10.2307/763685>.

———. *The Music Treatise of Anonymous IV: A New Translation*. Neuhausen-Stuttgart: American Institute of Musicology, 1985.

———. *Music in Medieval Europe*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

———. "Riches of Organum." *Early Music* 33 (4) (2005) : 708–710 <https://doi.org/10.1093/em/cah165>.

Žižka, Jan, František Dařena, and Arnošt Svoboda. *Text Mining with Machine Learning: Principles and Techniques*. London: CRC Press, 2020 <https://doi.org/10.1201/9780429469275>.