The Encoded Codex

A Digital, Statistical, and Codicological Examination of Paris, Bibliothèque nationale de France, fonds français, 24432

Sebastian H Dows-Miller
St Hilda's College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2024
97,440 words

Short abstract

The main objective of this thesis is methodological, showing the value of using methods from the digital humanities alongside analogue approaches to expand our understanding of Western medieval manuscripts and their texts. I take as a case study a manuscript collection of Old and Middle French texts (Paris, Bibliothèque nationale de France, fonds français, 24432), which has been puzzlingly neglected by scholarship.

The thesis' three chapters illustrate the different ways that digital methods can supplement more traditional methods, be that as the primary direction from which to explore a given feature, as a secondary mode of analysis, or as an equal partner alongside existing methods.

The first chapter presents a full codicological and palaeographical analysis of the codex. Within it, I use quantitative techniques in an equal partnership with more conventional qualitative methods. Drawing from both perspectives, I reach new conclusions regarding the manuscript's production, and the individuals involved in its creation. The second chapter then takes a digital-first approach, examining the use of scribal abbreviation, a hugely understudied area for French in particular, making this one of the major methodological contributions of the thesis. The third chapter presents a case study showing how methods of digital and statistical analysis can be combined with existing approaches on questions of uncertain authorship.

Throughout these chapters, I add to the toolbox of codicologists, palaeographers, and literary scholars, showing the utility of my methods beyond the scope of the thesis, while also making an important contribution to our understanding of this one particular manuscript.

Long abstract

The main objective of this project is methodological, demonstrating the value of combining methods from the digital humanities with more analogue approaches, to expand our understanding of Western medieval manuscripts and their texts. I take as a case study a fourteenth-century manuscript collection of diverse texts in Old and Middle French (Paris, Bibliothèque nationale de France, fonds français, 24432), which contains 446 leaves and nearly 67,000 lines of text, and has been puzzlingly neglected by modern scholarship.

As part of the foundational work for the thesis project, I completed a full manual transcription and encoding of the manuscript's contents in TEI P5-compliant XML. This is a contribution in itself, as it is already available in Open Access online for future scholarship to exploit, but its main function is to provide the core dataset at the heart of the digital, statistical, and codicological analyses contained across the chapters of the thesis.

The thesis' three chapters illustrate the different ways in which digital methods can supplement more traditional methods, be that as the primary direction through which a given feature should be explored, as a secondary mode of analysis, or as an equal partner alongside existing methods.

The first chapter takes the latter approach, presenting a full material, structural and palaeographical analysis of the codex. Within it, quantitative techniques are brought to bear on the encoded manuscript in an equal partnership with more conventional qualitative methods. We begin with an examination of the manuscript's structural complexity, using both qualitative and quantitative methods to distinguish its scribal hands, before turning to detailed examination of the manuscript's quire structure, the surviving evidence of pricking and ruling, the near-contemporary foliation, the *litterae notabiliores*, and the table of contents. Following this, I then use this evidence of complexity to produce an argument for the

sequence in which fr24432 was initially created and then subsequently recompiled, all of which took place, I argue, before it left its original place of production. We then examine what little can be gleaned from the surviving evidence about the manuscript's place of origin, early and later ownership, as well as the dates around which it may have been produced.

In appendices to Chapter 1, I go some way beyond the arguments presented in the body of the chapter, in order to contribute to broader scholarly debate about the codex, through further observations surrounding a number of notable codicological, paratextual, and textual features.

Drawing from both traditional and digital perspectives, in the first chapter I am therefore able to reach new conclusions regarding the manuscript's production, and the individuals involved in its creation, including the new discovery that no fewer than eleven scribes contributed to fr24432, a surprisingly high number for a manuscript of this kind. The mixed-method approach also delivers insights into the manuscript's development across time, as it passed between owners and readers, many of whom made some contribution to the material object we see today.

Following the opening analysis, the second chapter of the thesis takes a digital-first approach. Here, I present a case study for an examination of scribal abbreviation, an area of palaeographic and codicological research which has been hugely understudied for French in particular, and to which the digital humanities have a great deal to offer. In keeping with my mixed-method approach, I begin this chapter by taking a more theoretical perspective to abbreviation, considering its place within medieval writing practices, and the previous attempts that have been made by scholarship to account for its use. Through this work, I settle on an approach to manuscript abbreviation that is more concerned with function than form, arguing that this allows us to better account for the realities of how medieval scribes

used abbreviation. Having established the theoretical basis for our analysis, I then turn to methodological considerations, arguing again that the methods and approaches used by previous generations of scholarship regarding scribal abbreviation are not the most useful lenses through which to examine such questions. Instead, I propose a method based within the statistical coding language R, which makes use of digital and statistical techniques to identify trends and shifts in the use of abbreviation within a given manuscript. Applying this method to fr24432, I am able to reinforce conclusions about the scribal structure of fr24432 that are made in the preceding chapter, while also considering non-scribal factors which may impact the use of abbreviation in a given manuscript, such as the position of a candidate for abbreviation relative to other codicological features such as the beginning or end of a line, column, page, or quire. This examination of abbreviation in fr24432, while a significant contribution in and of itself to our understanding of the codex and a useful source of data surrounding fourteenth-century abbreviatory practices, is therefore also one of the major methodological contributions of the thesis more broadly. The approach I have developed to identify moments of scribal juncture can be replicated by codicologists and palaeographers when examining other manuscripts to help them identify when the work of one scribe ends and the work of another begins, thereby allowing them to understand better the circumstances of a given manuscript's production.

The third and final chapter of the thesis presents a case study displaying how methods of digital and statistical analysis can be combined with existing approaches on questions of uncertain authorship. Taking one example from within fr24432, I examine a corpus of texts found mostly within the codex which have often been attributed by modern scholarship to one Jehan de Saint-Quentin. I argue that the application of digital and statistical methods fits usefully alongside the more traditional methods of linguistic and textual analysis followed by previous scholars, allowing us to affirm the attribution of these texts to Jehan through

concrete, reproducible results, as well as permitting the addition of two further texts to the *œuvre* of the same author, namely *Le dit de Guillaume d'Engleterre* and *Le dit de Robert le Deable*.

In drawing my conclusions I examine the texts of this target corpus collectively and as a unit, as well as through comparison to a number of reference corpora and sub-corpora, ranging from those which share codicological contexts with the target corpus to those which exist as a distinct generic or authorial group, as well as much broader corpora combining texts with a variety of spatial and temporal origins. The features that I consider include more formal ones such as versification and shared vocabulary and structures, more textual features including details of claimed performance, genre, and sources, as well as more linguistic features. Following the completion of this analysis I am then ultimately able to conclude, from the holistic consideration of this evidence, that the texts of the reference corpus, as well as our two hypothesised additions, can comfortably be attributed to a single author. Again, further contributions to the literature surrounding Jehan de Saint-Quentin are made in an appendix to Chapter 3, in which I examine the codicological contexts of his works, focussing initially on their inclusion within fr24432, before moving to an examination of the contexts of those poems which appear elsewhere, to contribute to a broader understanding of this otherwise obscure poet.

As with the second chapter's contributions to palaeography and codicology, the third chapter therefore goes beyond the conclusions specific to fr24432, and further adds to the literary scholar's toolbox, demonstrating the utility of my methods to contexts far beyond the scope of the thesis, and reshaping the ways we think about authorship in the Middle Ages.

Table of contents

Introduction 1
The project 1
The encoding 3
1. The codex 6
1.1 Structural complexity 7
1.1.1 Qualitative scribal analysis 7
1.1.2 Quantitative scribal analysis 21
1.1.3 Quire structure 24
1.1.4 Pricking and ruling 32
1.1.5 Foliation 37
1.1.6 Litterae notabiliores 40
1.1.7 The table of contents 47
1.2 Production and recompilation 50
1.3 Localisation 53
1.4 Ownership 59
1.5 Dating 64
2. Sc^{i}bal abb'viatiō 68
2.1 The study of medieval French manuscript abbreviation 69
2.2 Theoretical considerations 70
2.2.1 Formal typologies 70
2.2.2 Functional typology 74
2.2.3 Processes of use and interpretation 79
2.3 Methodological considerations 81
2.3.1 Changing the object of analysis 81
2.3.2 Methods of analysis 83
2.4 Abbreviation in fr24432 84
2.4.1 Macro-level analysis 85
2.4.2 Micro-level analysis 88
2.4.3 Factors affecting abbreviatory usage 113
2.5 Conclusion 129
3. A la recherche d'un auteur perdu 131
3.1 Introduction 132
3.1.1 Jehan de Saint-Quentin 134
3.1.2 Existing attribution to Jehan de Saint-Quentin 135
3.2 Methodology 139
3.2.1 Use of statistical methods 140
3.2.2 Target corpus 141
3.2.3 Reference corpora 146
3.3 Form 149
3.3.1 Text length 151
3.3.2 Prologues and epilogues 155
3.4 Genre 158
3.4.1 miracle 159
3.4.2 exemple 164
3.4.3 dit 170
3.4.4 Generic descriptors outside of the J-corpus 172
3.4.5 Generic descriptors in rubrics 174
3.5 Performance 176
3.5.1 Evidence of orality 176
3.5.2 bone gent 177
3.5.3 seigneur 182
3.5.4 tous et toutes and ceulz et celles 186

3.6 Versification ... 189
3.6.1 Masculine and feminine rhymes ... 190
3.6.2 Rhyme frequency and 'easy' rhymes ... 198
3.6.3 Word frequency in rhyming position ... 205
3.6.4 Admixtures ... 212
3.7 Shared vocabulary and structures ... 216
3.7.1 Existing observations ... 216
3.7.2 Line-initial words ... 219
3.7.3 Onset of speech ... 229
3.8 Language ... 242
3.8.1 Localisation ... 242
3.8.2 Dating ... 244
3.8.3 Possessive adjectives ... 245
3.9 Sources ... 253
3.9.1 Claimed oral sources ... 254
3.9.2 Claimed written sources ... 256
3.9.3 A possible contradiction ... 265
3.10 Conclusion ... 267
Conclusion ... 272
Bibliography ... 276
Manuscripts ... 276
Printed works ... 276
Primary sources ... 276
Secondary sources ... 279
Digital resources and software ... 282
Appendices ... 284
Appendix A: fr24432 contents ... 285
Appendix 1A: Data ... 288
Appendix 1B: Further codicological observations ... 289
1B.1 Damage and repairs ... 289
Appendix 1C: Further paratextual observations ... 292
1C.1 Decoration ... 292
1C.2 Rubrication ... 297
1C.3 Explicit markers ... 301
1C.4 Textual layout ... 302
Appendix 1D: Further textual observations ... 305
1D.1 Form ... 305
1D.2 Versification and rhyme ... 306
Appendix 3A: F-corpus contents ... 311
Appendix 3B: Data ... 312
Appendix 3C: Scripts ... 317
Appendix 3D: Codicological analysis ... 318
3D.1 Manuscripts ... 318
3D.2 Jehan's texts in fr24432 ... 320
3D.3 Jehan's texts in other manuscripts ... 331
Acknowledgements ... 349

Introduction

The project

Paris, Bibliothèque nationale de France, fonds français, 24432 (henceforth fr24432) is a fourteenth-century codex written almost exclusively in Old and Middle French. It consists of 446 leaves, divided into two columns per side. The codex contains some 90 items in verse and prose, including a few additions by the manuscript's later readers, as well as a table of contents.[1]

The manuscript appears within well-studied groupings of codices, as manuscript *P* of the *fabliaux*,[2] manuscript *Q* of the works of Marie de France,[3] and manuscript *D* of the works of Rutebeuf.[4] It is, however, peripheral to the former two groups in particular, since it contains only one *fabliau* and one text by Marie de France, and, being from the middle of the fourteenth century, is quite late in the transmission cycle of all three.

The manuscript's less well known texts have been the focus of a few modern editions, particularly in the nineteenth century,[5] but there are still some which remain unedited, and most have received scant critical attention from modern scholarship. It is this which has been the primary motivating factor behind the work on this manuscript: the hitherto largely untapped richness of the material it contains offers scope for a near *tabula rasa* approach, allowing the methodological value of a mixed-method examination to come to the fore.

The main intended contribution of this project is therefore a methodological one: to demonstrate the value of combining methods from the digital humanities with more analogue

[1] For an overview of these items, and the short titles by which I will refer to the texts throughout the thesis, see Appendix A: fr24432 contents on p. 285.
[2] See, for example, Noomen & van den Boogaard (1983–1998).
[3] On which, see Busby (2011).
[4] Zink (1989–1990).
[5] The work of Achille Jubinal (1835, 1839–1842, 1846) is particularly notable.

approaches, in order to expand our understanding of Western medieval manuscripts and their texts. Through the three case studies presented in the following chapters, I examine the different ways in which digital methods can supplement more traditional methods, be that as a secondary mode of analysis, as in the codicological and palaeographical examination presented in the first chapter, the primary direction through which a given feature should be explored, as in our examination of scribal abbreviation in the second chapter, or as an equal partner alongside existing methods, as in our consideration of a case of uncertain authorship in the third chapter.

After a brief excursus into the principles behind the encoding of this codex, the three chapters of this thesis will thus proceed as discrete sections, given that despite focussing on the same manuscript, they are concerned with disparate topics and make use of distinct methods. When taken as a whole, however, the chapters combine to give new insights into our manuscript of focus, while also overlapping in their presentation of innovative methods with applications far outside the study of medieval French literature. Their coincidence demonstrates the immense value offered by mixed-method approaches to the field of manuscript studies, suggesting that the discipline is best served by an approach that works across the divide between digital and non-digital scholarship, making use of whichever approaches best serve the demands of a given research question.

The encoding

In order for any digital examination of the codex to be possible, it has first been necessary to complete a full encoding of the codex within P5-compliant TEI-XML. This encoding, which has been deposited in an open access repository,[6] has been completed primarily with reference to an online facsimile of the manuscript itself,[7] but also to previous editions of the manuscript's texts, where they exist, as well as in-person consultation, and serves as the basis for much of the avenues of analysis that I pursue throughout the thesis. The completed encoding does not follow a single transcription approach, rather it presents a combination of more diplomatic approaches, retaining variant letter forms and original punctuation, alongside regularised versions of the text, in which letter forms and punctuation markers and are standardised to their modern equivalents. This allows for a variety of transcription types to be extracted from the same encoding, depending on the requirements of the analysis at hand.

Individual texts, and their *incipit* and *explicit* markings (connected to the main text through `@corresp`) have been separated through use of the `<div>` tag, with sections of prose marked using `<p>`, and sections of poetry marked using `<lg>`, with individual lines of verse tagged using `<l>`. When a text is dialogic in form, this is marked using `<spGrp>`, with contributions by individual speakers marked using `<sp>`. Direct speech within non-dialogic text has been marked with `<q>` where such details are pertinent to our analysis. Both `<sp>` and `<q>` make use of `@who` and `@toWhom`, where appropriate, and instances of `<q>` that break across material divisions are connected through the use of `@next` and `@prev` references. The few words and phrases written in languages other than Old and Middle French (predominantly Latin) have

[6] Available at https://zenodo.org/doi/10.5281/zenodo.12608780.
[7] Available at https://gallica.bnf.fr/ark:/12148/btv1b52516597g.

been marked using @xml:lang and contained within <seg>, if these do not align with another tagged section of text.

Material divisions within the codex have been marked, with <gb>, <pb>, <cb>, and <lb> tags indicating the beginning of a new quire, folio, column, and line respectively, with the latter tag marked with @break as "no" to indicate a line broken within a word. <handShift> is used to mark the beginning of a section of work by a particular scribe, with the scribe identified using @scribeRef. Additions and interventions by hands other than the main scribe of a section are marked through the addition of the @hand attribute.

The table of contents is encoded as a table using the <table>, <cell>, and <row> tags. Labels, marginal notes, catchwords, and highlighted sections are marked using <label>, <note>, <fw>, and <hi>, with information about the location and presentation of these additions usually included in @place and @rend attributes, the latter of which is applied to material and textual divisions more broadly, particularly to mark the use of *litterae notabiliores*, or to indicate the exact manor in which text is added or deleted. Meanwhile, moments of damage, blank spaces, and gaps in text are marked using <damage>, <space>, and <gap> respectively. In the latter case, details such as the @agent, @extent, @quantity, and @reason behind the gap are marked. All these material and textual divisions are given a unique identifier through the use of @xml:id.

Within the text of the encoding, expansions of scribal abbreviation, additions and deletions by scribal or later hands, as well as editorial regularisations, corrections, suppressions, and insertions have been encoded using the <expan>, <add>, <del>, <reg>, <corr>, <surplus>, and <supplied> elements respectively, alongside the unexpanded, unregularised, and uncorrected version of the text, through the <abbr>, <orig>, and <sic> elements respectively. Such modifications are marked with @resp to indicate the responsible party, where this is

known. Where more than one variant of the same segment of text has been encoded, for example in cases of expanded abbreviation, correction, and regularisation, the variants have been contained within a `<choice>` tag.

Again, where such details have been pertinent to our analysis, personal names, geographic names have been encoded within `<persName>` and `<placeName>`, each referring to an associated identifier using `@ref`. Dates are encoded in `<date>`, usually with an `@when` attribute, and `<num>` elements are encoded with their associated `@value`, to enable machine reading of complex numeral sequences.

The encoding of fr24432 therefore seeks to include all those details which are relevant to the analysis presented in this thesis, but also to serve as a broader source of data for future scholarship to exploit, including as a model for projects approaching the full encoding of a single manuscript. There are many further avenues that could be pursued from these data which we will not pursue here, most notably the examination of references to people and places, as well as scribal uses of different kinds of modification. We now turn to our initial codicological examination of fr24432, in which these data play a significant part, although they will be most pertinent to the analysis in the second and third chapters.

1. The codex

A codicological and palaeographic study of Paris, Bibliothèque nationale de France, fonds français, 24432

This chapter considers fr24432 as a material object, examining its form, structure, and history, while also asking what the codex can tell us about the parties involved in its creation. The principal thesis of the chapter is that fr24432, far from being a homogenous unit conceived and completed as a single codex, is in fact a composite manuscript that underwent a period of evolution, modification, and change, most of which occurred early in its life. The codex has been in something resembling its current form since before it left the *scriptorium* in which it was created, and yet at the same time we can see that this process of creation did not follow a linear path.

The evidence for this, as we shall discover, comes from the fact that the manuscript displays a great deal of diversity in certain features predating its completion, and a high degree of unity in others.[1] In doing so, we will draw from traditional palaeographical and codicological methods, as well as through emerging, data-driven techniques. The culmination of this evidence will ultimately lead us to conclude that the single manuscript fr24432 was originally at least three distinct manuscript units. Following this, we will then turn our attention to broader questions surrounding the manuscript's localisation, early ownership, and dating.

1.1 Structural complexity

1.1.1 Qualitative scribal analysis

Let us begin our examination of the internal diversity of fr24432 by taking a qualitative, palaeographical approach to its scribal hands. The analysis here, the first detailed considerations of the hands within fr24432, suggests that there are a total of eleven separate scribal hands involved in the main body of the codex.[2] All fit broadly into Derolez's (2003:

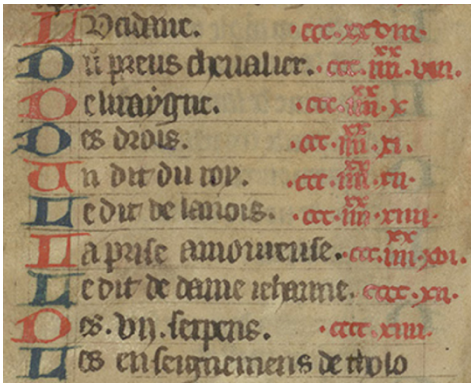
[1] There is much that can be said about this codex that does not point in either of these directions and therefore is beyond the scope of our overall argument, but which nevertheless offers valuable contributions to our understanding of this manuscript. A number of observations of this type are included in appendices below.
[2] All numbering used here is my own, based arbitrarily on the order they first appear in the manuscript in its current form. Previous analyses have only suggested the presence of a few hands, for example Mercier & Thomas (1894: 465) claim the codex is 'écrit tout entier de la même main,' while Grange (2018) suggests there are 'probably two or three similar hands.' Olaf Posmyk (2013) has suggested that 'l'idée selon laquelle le

72–101) description of *Northern Textualis* script, but with a good deal of variation between them.

Distribution

Hand 1 is found only in the table of contents (ff. 0ra–0va), and writes in a black ink which has faded in places. It will be argued below that the table was likely added by the same *scriptorium* that produced the main body of the manuscript, but at a slightly later date, and so it is unlikely to be exactly contemporary with all the other hands in the manuscript.

Figure 1.1 – sample of Hand 1 (f. 0ra).

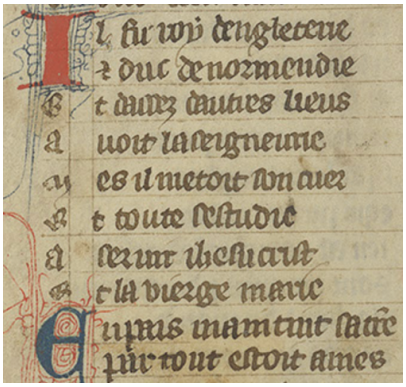
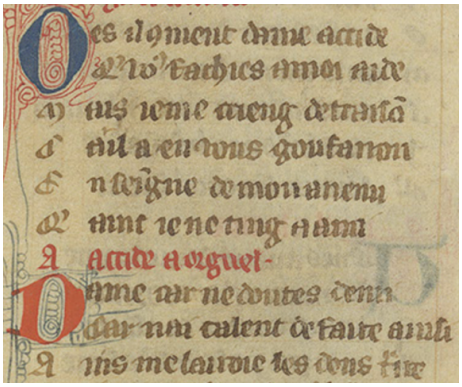


Hand 2 appears at both extremes of the manuscript in its current ordering. It is used for four separate stints (ff. 1ra–29vb l. 15; 33ra–56vb; 388ra–416va l. 9; all of f. 419), but it is more productive to see these as two runs (ff. 1ra–56vb; 388ra–419vb), temporarily interrupted in both cases mid-column by another hand (hand 3 in the former case and hand 11 in the latter), after which hand 2 returns to complete the run. The fact that both interruptions begin in the same column (and in the case of hand 11's interruption, within a single word) suggests a high degree of coordination between hands 2, 3, and 11.[3]

manuscrit a été copié par un seul scribe est à écarter,' but did not go so far as to suggest how many scribes were involved in the its production.

[3] This apparent cooperation highlights a close association between the sections containing the work of hand 2 at the beginning and end of the codex, despite the presence of intervening stints by other hands, supporting the hypothesis articulated below that both sections formed part of an original 'base' manuscript unit, to which others were added.

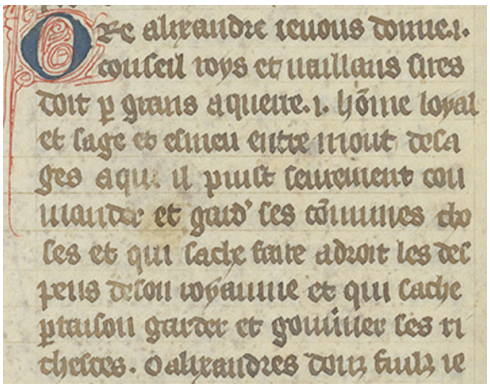
It is of note that hand 2 is only used in stints containing verse texts, whereas hand 11, after its brief interruption into *Miroir*, a verse text, then copies two prose texts in its much longer second stint, perhaps suggesting a predilection amongst certain scribes for copying texts of particular types. The final text of the manuscript, *Misere*, which appears at the end of hand 11's final stint, is incomplete, and so it is also possible that this section, and hand 11's contribution, was once much longer than is currently visible. Hand 2 predominantly writes in a dark brown ink, sometimes close to black, whereas hand 11 writes in a lighter brown ink.

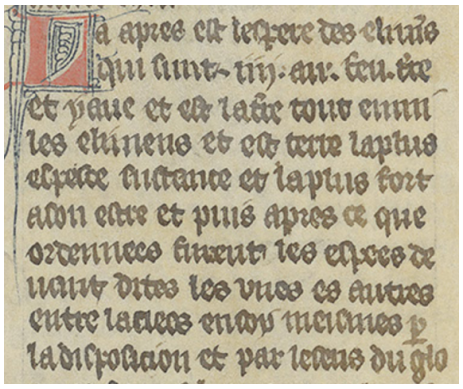
Figure 1.2 – left: sample of Hand 2 (f. 1va); right: sample of Hand 11 (f. 417rb).



Hand 3, after its brief interruption of hand 2's first run (ff. 29vb l. 29 –32vb), completes a further three stints (ff. 57ra–72vb; 76ra l. 10 –94vb; 239ra–246vb), although as with hand 2 the first two of these appear to be a single run interrupted mid-quire by another hand.[4] The first two stints are completed in a light brown ink, and the latter two stints are completed in a darker ink, closer to black. The interruption between hand 3's first two stints is by hand 4, which appears only in this very short stint (ff. 73ra–76ra l. 10) and in a darker ink than the surrounding text in hand 3, with hand 3 returning in the middle of a line. Hand 3 is used across both verse and prose texts, and its final stint appears incomplete, given that the final text in the quire (*Trois mortes*) is fragmentary.

[4] We will examine the manuscript's scribal structure in further detail below, from p. 24.

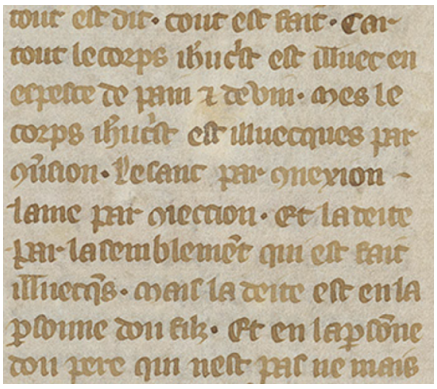
Figure 1.3 – left: sample of Hand 3 (f. 60^{vb}); right: sample of Hand 4 (f. 78^{vb}).

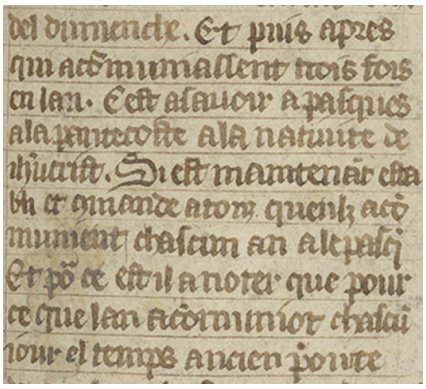




Hand 5's contribution consists of two stints totalling three complete quires (ff. 95ra–102^{vb}; 311ra–326^{vb}), containing both prose and verse texts. In both cases, these stints are not discrete, and instead the texts contained within them overlap with other hands in adjacent quires. In the first stint, this is the case at both the beginning (hand 3, *Vision*) and the end (hand 6, *Vie du monde*), whereas in hand 5's second stint, the overlap only occurs at the end of the text in question (*Messe*), where hand 9 completes the final two columns.[5] Hand 5 writes in a light brown ink, whereas the contribution by hand 9 is in a darker ink.

Figure 1.4 – left: sample of Hand 5 (f. 323^{vb}); right: sample of Hand 9 (f. 327ra).



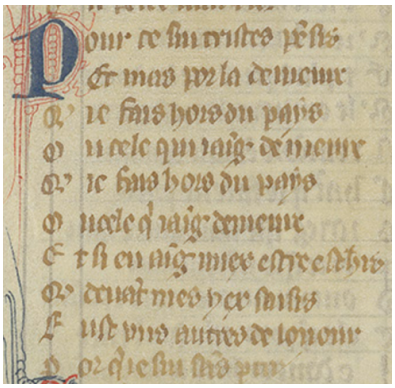


Hand 6 completes one single, long stint (ff. 103ra–198G^{vb}), which at 404 columns makes it the most prolific hand in the manuscript.[6] The scribe writes in a light brown ink and with a

[5] This is the only text copied by this hand, and may serve to resolve a lacuna resulting from a later process of recompilation, which we discuss further below.
[6] This stint includes folios 198A–G, a single quire labelled later as they were missed from the original foliation.

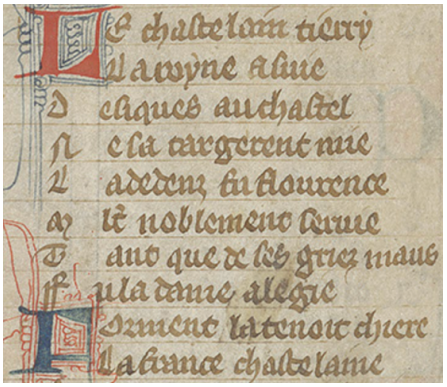
visibly condensed script. The hand does not appear in any quires containing other hands, but the stint begins within *Vie du monde*, which is shared with hand 5, and the final text in the stint (*Lai d'amour 1*) is fragmentary, so it is possible that further cooperation with other scribes appeared in whatever following material has been lost.

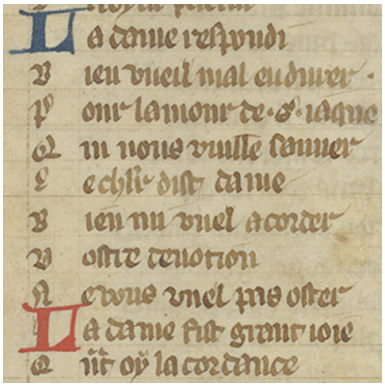
Figure 1.5 – sample of Hand 6 (f. 169ra).



Hand 7 appears in two stints to copy short verse texts of varying metres (ff. 199ra–230vb; 247ra–310vb), with intervening quires in hand 8 (ff. 231ra–238vb; the only instance of this hand in the manuscript), and hand 3 (239ra–246vb; as discussed above, this was possibly the first quire of several). Both hands 7 and 8 write in a similar, mid-brown ink. In both cases hand 7's stint is discrete in terms of both texts and quires, and as such it is conceivable that what now appear to be two stints once formed part of a single whole.

Figure 1.6 – left: sample of Hand 7 (f. 221rb); right: sample of Hand 8 (f. 232rb).





Hand 10's only contribution is the prose text *Lucidaires*, and is discrete in its quires, with its final quire (51) cut short rather than continuing with the work of a new scribe. Hand 10 writes in a dark brown ink that sometimes tends towards black, and the discrete nature of the contribution gives us our first indication of the possibility, for which we will see further evidence in the following sections, that the work of hand 10, and with it *Lucidaires*, was added during a later process of recompilation.

Figure 1.7 – sample of Hand 10 (f. 329ra).

Minuscule letter forms

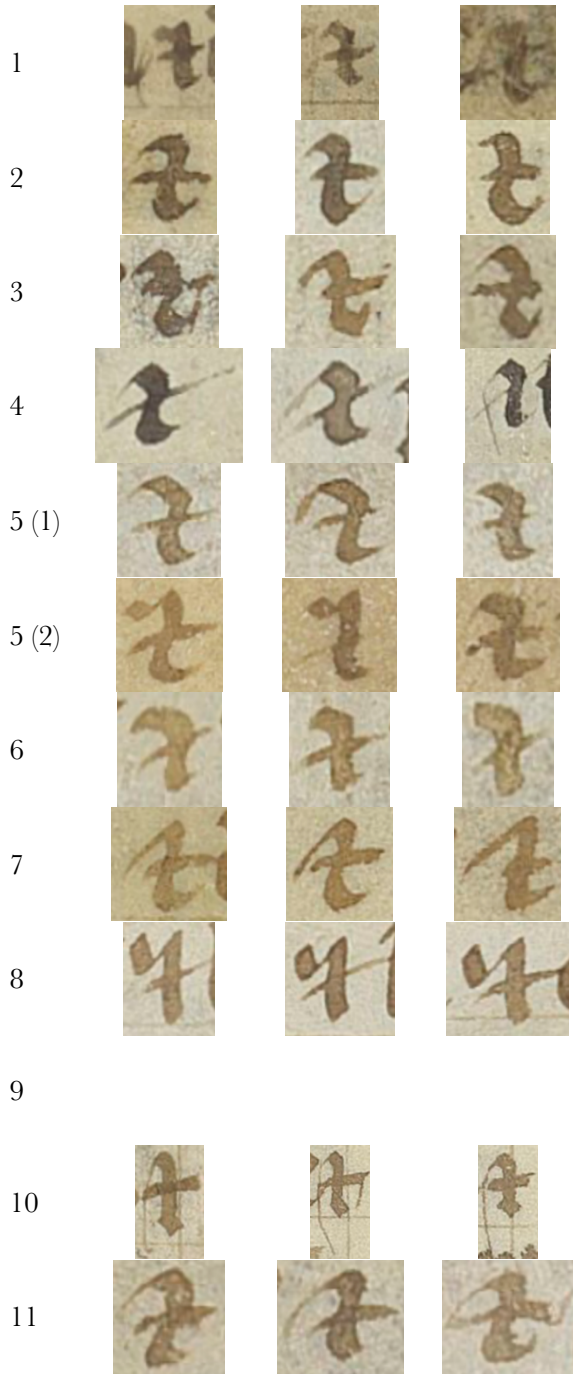
Given the widespread variation between majuscule characters even within stints by the same scribe, the most convincing qualitative evidence for the eleven distinct scribal hands comes from an examination of the manuscript's minuscule characters. On a general level, hands 1, 9, and 10 could be described as examples of *textualis formata*, given their somewhat more angular style, while the remaining hands fit better into the category of *textualis libraria*, although hands 6, 8, and 11 arguably tend a little towards the *textualis currens*, given their condensed writing style. Even given these divisions, the characters of one scribe often share similar features with those of another; it is the combination of these features across all characters that points to the presence of a unique hand.

⁊7⁊

The stem of the Tironian note has a markedly angular feel in hands 1, 8, and 10, and is somewhat angular in hands 3, 4, and 5. In hands 2, 6, 7, and 11, the form of the stem is more arced. No Tironian notes appear in either of the columns of hand 9.[7]

Hands 4, 7, 8, 10, and 11 display a swash extending down from the top of the stem, sometimes extending to or beyond the full height of the character (4, 7, 10), and sometimes roughly half the character's height (8, 11). The swash for hand 10 descends in a largely vertical direction, whereas in the other hands it is closer to the diagonal, although on occasion this is matched in hand 4. In hand 8 the swash is terminated by a near-vertical full width stroke.

Table 1.1 – examples of the Tironian note.

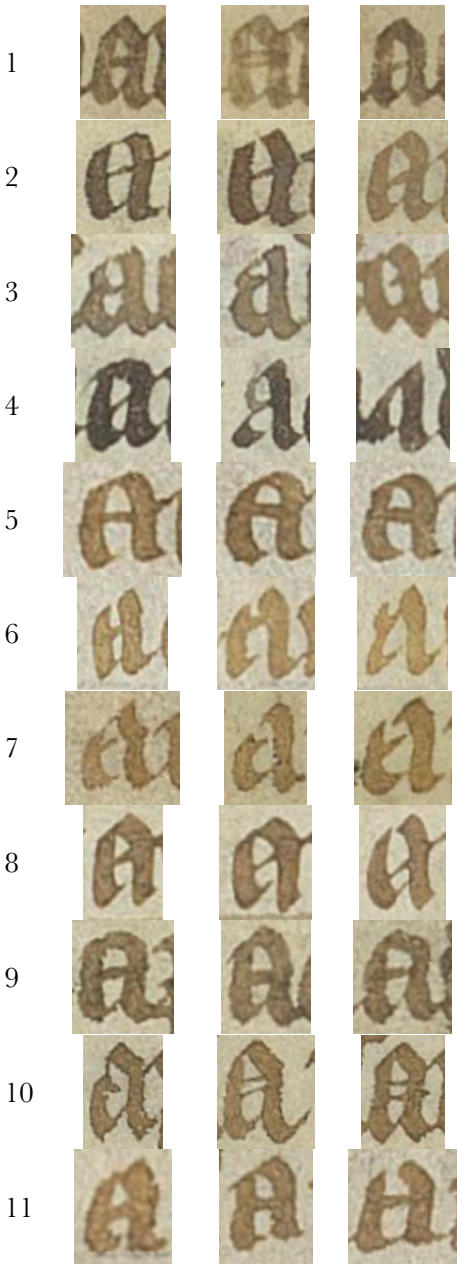


Hand 5 displays formal variation between its stints (ff. 95ra–102vb; 311ra–326vb), with the stem usually topped by a jointed arm in the second stint, but generally more rounded in the first.

[7] The examples given on these pages have been selected in a somewhat random way, but examples have been privileged that allow a clear view of the whole character. In cases where a character engages in biting or fusion with a following character, I include such an instance as the third example for a given hand.

⟨a⟩

Table 1.2 – examples of the character ⟨a⟩.



In almost all cases the manuscript's hands prefer to present a two-compartment ⟨a⟩, with the two counters closed by a single horizontal stroke, although some variation does exist.

Hands 1, 2, 5, 7, 8, 9, and 11 are relatively consistent in their preferences. For hands 1, 2, and 9, the top of the bowl is formed by leftmost ascender, while in hands 5, 7, 8, and 11 the leftmost stroke is shorter and it is instead the rightmost stroke which closes the aperture.

The crossbar which divides the character into two compartments is roughly horizontal in hands 1, 5, 9, and 11, but tends diagonally up right in hands 2, 7, and 8.

Hand 6 is unusual amongst the hands of the manuscript in that the left side of the ⟨a⟩ usually consists of two loops, of which the uppermost is not always fully closed.

Hands 3 and 4, meanwhile, are particularly variable. Both display examples of the structure common for hands 1, 2, and 9, with two vertical strokes closed by a shoulder in the leftmost stroke, and split by a crossbar. Both also display the method preferred by hand 6, with the two counters closed on the left by two separate strokes. Hand 4 in particular also displays use of a single-compartment ⟨a⟩, consisting of a single counter with no crossbar.

Hand 10 has different forms for word-initial and word-internal instances of ⟨a⟩. The word-initial form is somewhat like that preferred by hand 6, with the left side of the character created through the use of two loops. The word-internal character resembles that used by hands 1, 2, and 9, although in this case the two counters are separated by two adjacent crossbars, rather than one, which is a trend also seen occasionally in the work of hand 1.

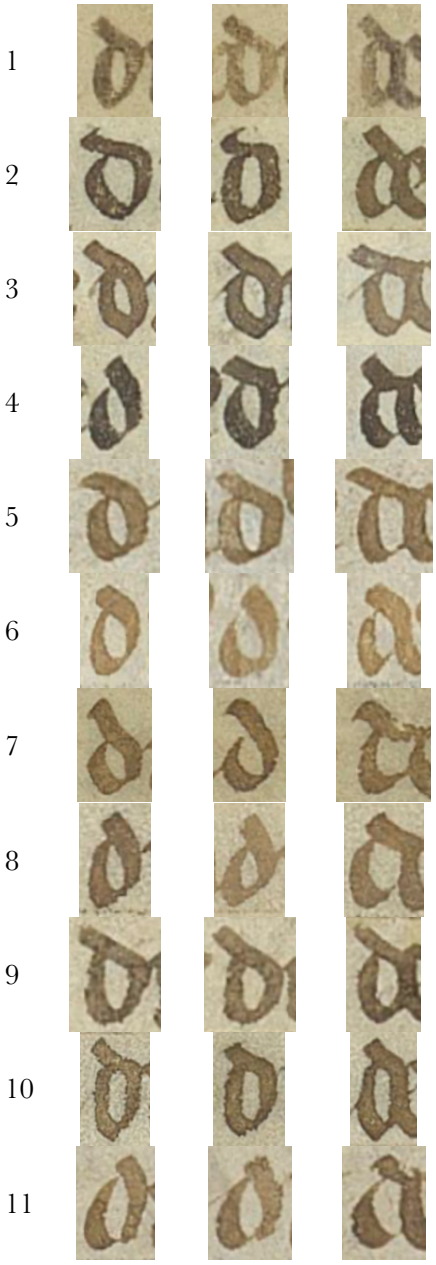
⟨d⟩

The ascender of ⟨d⟩ varies in height, even within the same hand, although there are some general trends with, for example, hands 3 and 7 tending towards longer ascenders, and hands 6 and 11 towards shorter ones.

In hands 5 and 7, and to a lesser extent also hand 3, the ascender stroke is terminated by a movement of the quill diagonally down left, whereas in all other hands the trend is for termination through a movement of the quill diagonally up right.

In all hands besides 1, 9, and 10, ⟨d⟩ consistently engages in biting with a following rounded character through the non-completion of the stroke containing the ascender, allowing the bowl's aperture to be closed by the character that follows the ⟨d⟩. In hands 1, 9, and 10, this combination is achieved simply through the close juxtaposition of the two characters.

Table 1.3 – examples of the character ⟨d⟩.



⟨b⟩

Table 1.4 – examples of the character ⟨b⟩.

The ascender of the character ⟨b⟩ exists in two different variants, with hands frequently containing both but usually preferring one over the other. Hands 5, 8, and 11 prefer a first variant, namely the inclusion of a hairline approach stroke that extends leftwards or diagonally down left. The alternative variant consists of a beak that extends diagonally up left, giving the impression of a bifurcation of the ascender. Sometimes this is more pronounced, as in the work of hands 2, 3, 4, and 7, and in others it is less so, as in hands 1, 6, 9, and 10.

⟨b⟩ varies more than ⟨d⟩ in its combination with following rounded letters. Hands 3, 4, and 11 tend to prefer non-completion of the rightmost stroke, as in ⟨d⟩, whereas hands 1, 9, and 10 favour close juxtaposition of two characters without biting. The other hands prefer something between both extremes.

⟨h⟩

The ascender of the character ⟨h⟩ closely matches the hand's preferred variant of the character ⟨b⟩: hands 5, 8, and 11 continue to prefer the hairline beak extending down left, while the remaining hands prefer bifurcating embellishment of varying intensity. There does however appear to be greater variation in the embellishment of the ascender of ⟨h⟩ than in the case of ⟨b⟩.

The height of the ascender varies within hands, but hands 1 and 9 are notable in particular for their shorter ascenders.

There is a great deal of variation within hands in the extent to which the shoulder of the second stroke is curved or angular. That said, those in hands 1, 2, 3, and 10 in particular appear to tend towards a more angular shoulder with a more vertical descender, while in hands 4, 6, 8, and 11 both are usually markedly more curved.

In most cases, the second, descending stroke begins to taper towards its terminal at the character's baseline. Sometimes the taper begins below the baseline, as in hand 5, or above it, as in hand 4.

The terminal of the descending stroke is usually unembellished, but hand 2 and occasionally also hand 11 include a retroflex arc, caused by the quill being lifted fractionally late as it is returned above the baseline to begin the next character. The terminal of the descender in hand 4, meanwhile, most likely resulting from the same phenomenon, often resembles a ball terminal.

Table 1.5 – examples of the character ⟨h⟩.

As with the Tironian note, hand 5 displays variation between its two stints. It is not uncommon for the descending stroke in the first scribal stint to be straight and to continue below the baseline before tapering, whereas this is less common in the second stint, which

could suggest a time gap between the production of these stints, or that hand 5 was influenced by external factors such as the exemplar from which they copied.[8]

⟨g⟩

The rightmost vertical stroke in the main bowl of ⟨g⟩ is usually curved in all but hands 1 and 10, where it is straighter. The same is true of the leftmost vertical stroke, although here hands 6 and 9 are often straighter.

In hands 1, 2, 3, 5, 8, 9, and 10 the descender of the character ⟨g⟩ largely extends leftwards, with minimal curvature for some (1, 9, 10) and greater curvature for others (2, 3, 5, 8). In this latter group, a hairline stroke sometimes returns diagonally up right, closing the the counter. In hands 4, 6, 7, and 11, meanwhile, the descender usually forms a loop, resulting in the creation of a full bowl.

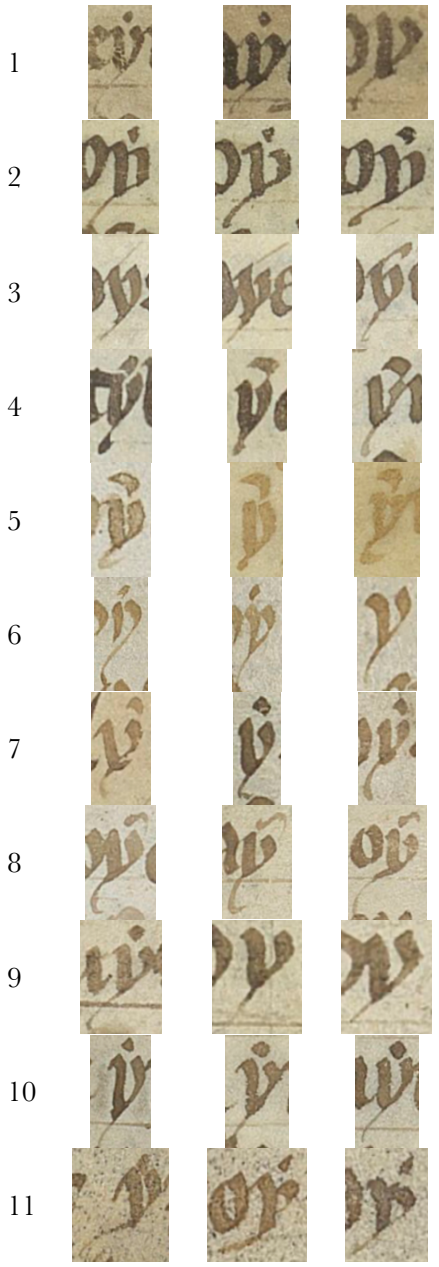
Examples of a horizontal ear used to connect a ⟨g⟩ to a following character can be found for all hands, however it is less common in hand 3.

Table 1.6 – examples of the character ⟨g⟩.

[8] Given the variation between the stints of hand 5 seen for both ⟨h⟩ and the Tironian note, one might reasonably question whether both stints were completed by the same scribe. On balance, the similarity of forms elsewhere, particularly in ⟨g⟩, ⟨y⟩, and ⟨z⟩, does appear to support assigning both stints to hand 5.

⟨y⟩

Table 1.7 – examples of the character ⟨y⟩.



The leftmost stroke of ⟨y⟩ is vertical with a rightward curvature in hands 1, 2, 6, 10 and 11, and with a lesser curvature in hands 4, 5, 8, and 9. In hand 7 it is frequently slanted leftwards, with hand 3 displaying both this slant and a leftwards curvature.

The rightmost stroke of the ⟨y⟩ is terminated by a descender, which in hands 1, 3, 7, 8, 9, 10, and 11 usually consists of a single diagonal hairline stroke, although in hands 6, 7, and 8 this is sometimes terminated by a retroflex arc which points to a greater or lesser extent back to the body of the character. This feature is also seen in hands 4 and 5, although it is much more pronounced in hand 5, and both hands differ from the others in that the descender retains a certain degree of thickness. Hand 2 varies in its presentation of the descender, which resembles a single hairline stroke with either no terminal, a retroflex arc, or a ball terminal.

⟨y⟩ is usually topped with a stroke of varying type, although in some cases across many hands it is absent. In hands 1, 2, 6, 9, 10 and 11 it resembles a *punctus*, whereas in hand 3 it is akin to the hairline stroke often found on ⟨i⟩. In hands 4, 5, and 7 it takes the form of a curved stroke descending diagonally rightwards, which resembles the curved shoulder of the character ⟨h⟩

found in all three hands. A similar stroke resembling a breve is the norm in hand 8, although on occasion this is inverted.

⟨z⟩

Table 1.8 – examples of the character ⟨z⟩.

The uppermost stroke of ⟨z⟩, which in all cases resembles the Arabic numeral ⟨3⟩, is largely comparable across all hands, but the second, lower stroke varies. In all hands but 3, 6, and 7, the initial sections of the stroke closely resemble the uppermost stroke. In hand 3 the stroke begins with a near-horizontal bar, followed by a diagonal descender. Hand 7 displays similar features, although the initial part of the stroke is closer to diagonal than horizontal. Hand 6, in keeping with its characterisation as closer to *textualis currens*, is more variable in the direction of its bottommost stroke, with the character often simply resembling a single vertical stroke that oscillates as it descends.

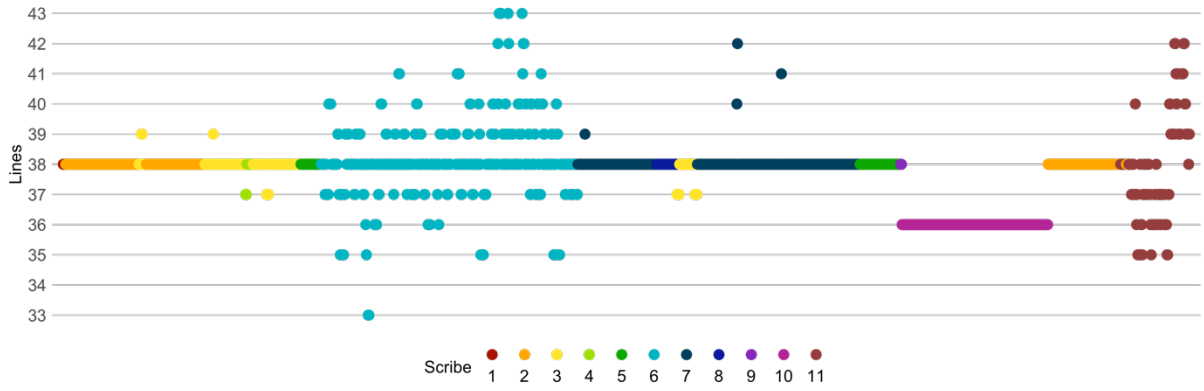
The descender of this bottommost stroke tapers to a diagonal hairline in all hands, although on occasion it is embellished with a retroflex arc, particularly in the work of hand 4. The descender varies in length across the hands, with hand 10 offering the longest stroke and hands 3 and 5 amongst the shortest.

1.1.2 Quantitative scribal analysis

It is therefore possible to conclude from the qualitative data that fr24432 is the work of 11 scribes, although in some cases these conclusions are tentative. Also taking a more quantitative approach can strengthen this suggestion by pointing towards the presence of different preferences across the scribal stints. In some cases, as we shall see, this variation is clearly a case of scribal preference, but in others the data tell a more complex story, suggesting that the scribes' exemplars may also have played a role in determining certain features of the end result.[9]

Quantitative analysis of the length of the manuscript's columns reveals a number of notable differences in scribal practice. The default column length is 38 lines, however we observe some variation across hands, as seen in Figure 1.8.[10]

Figure 1.8 – column length, by hand.



Hand 10 is the only hand in the manuscript to prefer 36-line columns. Given that the stint only contains one text (*Lucidaires*), it is unclear whether this is as a result of scribal or ancillary preference, or the structure of the scribe's exemplar, but at the very least this adds to a

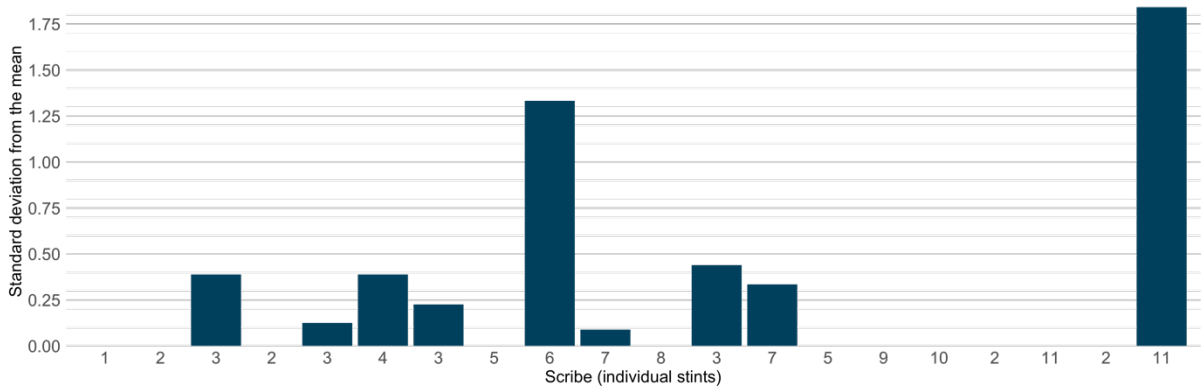
[9] There are a number of other sources of quantitative data which could be used to support the conclusions of this section, particularly the use of scribal abbreviation. Examination of this feature is however reserved for Chapter 2, which contains a much more comprehensive study than would be possible here.
[10] For the purposes of this analysis, in order to ensure consistency, column length is taken to include empty ruled lines, rather than just those which contain text.

growing sense from the data we have examined so far that the work of hand 10 is notably distinct from much of the rest of the manuscript.

The long stint of hand 6 is notable in its variability, despite containing texts similar to those found in the work of other scribes, even to the point of potentially being from the same exemplar.[11] While there is nevertheless a clear preference for columns of 38 lines, with 222 columns out of a total of 404 (55.0%) appearing at this length, the stint also contains both the longest and shortest columns in the manuscript, at 43 and 33 lines respectively, and has a particularly high rate of variability.

Figure 1.9 displays the standard variation of the columns in each scribal stint, that is, the average amount by which a given column would be expected to deviate from the default column length in that stint.[12] Hand 6 is indeed notable amongst these data, with a standard deviation of 1.34 orthographic lines, but this is hardly the only significant revelation.

Figure 1.9 – the standard deviation from a standardised mean per scribal stint, in orthographic lines.



[11] See, for example, the 26 moralising *dits* now attributed to Jehan de Saint-Quentin, of which 13 appear within the work of hand 6. We will discuss these in greater detail in Chapter 3.

[12] For the data behind this figure, see Table 1.46 on p. 288. Note that in calculating the standard deviation I have used a manually standardised mean, rather than the mean of the raw data, to reflect the reality that the default column length can only be an integer. This does not significantly affect the data since in all cases the mean column length per scribal stint is within 0.25 orthographic lines from the integer taken to be the default column length.

Hands 2 and 5 are consistent in their rigid adherence across all stints to columns of 38-lines, itself the most common format in the manuscript, which indicates that some scribes paid closer attention to the lengths of their columns than others, and thus that the influence of scribal preference is at play here.[13] Hands 1, 8, 9, and 10 are also fully consistent in their use of their preferred column length, although given that each hand only appears in one stint, it is unclear whether this is truly a case of scribal preference or due to some other factor, and in the case of hands 1 and 9 in particular, the stints are too short to be able to draw any concrete conclusions at all.

Hand 3 consistently displays a small amount of deviation across all four stints, a phenomenon also observed for hand 7, which suggests that these scribes were more willing to accept the odd moment of variation, although not necessarily to the extent seen for hand 6. Again, while it is possible that the variation within the single stint of hand 4 is due to scribal preference, rather than influence from the surrounding stints by hand 3, or from the exemplar, there is insufficient evidence to reach a conclusion either way.

Hand 11 is of particular interest. In its first, short stint, an interruption of a longer run by hand 2, no variation in column length is observed, whereas when the hand returns for its second and final stint, column length varies dramatically, displaying a standard deviation of 1.84, the highest of any scribal stint. The fixed column length found in the first stint of hand 11 can most likely be explained by the fact that the stint is only a subsection of quire 55, which starts and ends with the work of hand 2, and therefore all the folios of this quire were likely ruled at once, either by hand 2 or by some ancillary. Hand 11 was in this case likely just

[13] Evidence for our assumption here that, for the most part, scribes ruled their own quires (or at least oversaw their ruling) is explored in the section on pricking and ruling on p. 32.

abiding by the ruled lines assiduously set out by their colleague, with the variation occurring only when hand 11 gained overall responsibility for a given quire.

If we assume again that hand 11 ruled the quires of their second stint themself, or at least oversaw their ruling, then the variability seen in this stint therefore results from scribal preference, and a less rigid approach to laying out a quire's ruled lines. It is even possible that hand 11 felt no need for column length to be consistent between quires, since within this final stint, we observe a shift in how hand 11's columns deviate from the norm. In quires 56–57, the first two quires of the stint which contain most of *Salomon*, column length rarely reaches higher than 38, and has an average of 36.8. The next item, *Misere*, begins on the second folio of quire 58 (f. 437^{r}), in which a 38-line column appears to be the minimum, rather than the maximum, with an average column length of 39.9.

1.1.3 Quire structure

Examination of a number of scribal features present within fr24432 has helped to paint a picture of the manuscript's internal diversity. By taking a closer look at the materiality of the codex through its quiring structure, we can gain a further understanding of the extent to which scribal variation alone can explain this diversity, and to what extent other factors of recompilation and modification are at play.

Fr24432 contains a total of 59 quires, including a singleton incorporated at the beginning for the table of contents, with two folios displaced (ff. 255 and 262) and a number of quires missing. Quires 32 and 58 are clearly missing following quires, given that the items with which they end are fragmentary. For quires 6 and 26, given that the final leaf of each quire is missing, it is impossible to tell whether the missing material is limited to a single leaf or is more significant. In quires 7, 18, 30, 32, 34, 39, 40, 43, 51, and 55 the end of the quire corresponds with the end of an item, meaning that it is possible that following quires are

missing, but there is no evidence to suggest that this is the case. Meanwhile, a jump in the early foliation between qq. 51 and 52 from ff. 385 to 388 suggests that two folios may be missing from q. 51, which were potentially blank (like f. 385^{v}) and were therefore later removed as surplus, a hypothesis which is supported by the fact that the remaining leaves of q. 51 are bound as singletons.[14]

A formal structure, amending that given in Grange (2018), is therefore as follows:

> singleton (bound with 1), 1–5:8, 6:8 (wants 8), 7:8, 8:6, 9–10:8, 11:6, singleton, singleton, 12–24:8,[15] 25:6, 26:8 (wants 8), 27–29:8, 30:8 (wants 1), 31–32:8, wants quire(s), 33:8, 34:8 (1, 8 displaced),[16] 35–38:8, 39:10 (1 displaced from 34, booklet of 8, 8 displaced from 34), 40–42:8, 43:2 (bound as singletons), 44–50:8, 51:4 (bound as singletons; wants 3, 4), 52:8 (bound as: singleton, booklet of 6, singleton), 53–58:8, wants quire(s)

The folios of the manuscript measure around 310 x 215mm. The text block, which consists of two columns in all cases where there is sufficient text to fill the page, generally measures between 215 and 225mm in height, and between 140 and 150mm in width, although this latter measurement in particular varies between quires containing predominantly poetry and those containing prose.[17]

Grange (2018) divides the manuscript into three parts, with part 1 containing q. 1–25 as well as q. 44–58, part 2 containing q. 26–43, and part 3 containing q. 44–51. Close examination of the manuscript's texts and materiality, however, reveals that the reality is somewhat more complicated. If we examine groupings of quires which are discrete from other groupings at

---

[14] For more on the manuscript's foliation, see below.
[15] Grange (2018) suggests that quire 12 contains ten folios respectively, but it is clear from an examination of the manuscript that the first two leaves (87 and 88) which Grange includes in quire 12 are in fact singletons, included at the end of quire 11 to provide a full set of 8 leaves, as evidenced by the presence of catchwords at the end of f. 88^{v}.
[16] This displacement, caused by the removal of the outer bifolium from q. 34 and its binding with q. 39, appears to have been a much later modification, therefore my analysis of the manuscript will proceed as if this bifolium were in its rightful position.
[17] The text block in quires 37, 38, and 40, for example, is narrower than elsewhere, at around 135mm, which seemingly correlates with the presence of texts which are written in prose or a prose-like format.

the beginning and end, in that they do not contain overlapping texts, we find that there are twelve discrete sections, as represented in Table 1.9.

Table 1.9 – groupings of quires in fr24432 by discrete textual content.

Grouping	Quire(s)	Scribal hand(s)	Item(s)	Final catchwords hand(s)
a	0	1	Table of contents	–
b	1–7	2, 3	1–20	α
c	8–26	3, 4, 5, 6	21–52	β
d	27–30	7	53–56	γ
e	31–32	8, 3	57–60	3, α
f	33–34	7	61–63	β
g	35–39	7	64–71	β
h	40	7	72	β
i	41–43	5, 9	73–77	–
j	44–51	10	78	–
k	52–55	2, 11	79–88	β
l	56–58	11	89–90	11

With the exception of section ***l***, which is fragmentary, all catchwords present at the end of these sections are completed by non-scribal hands (Figure 1.10), named here as α, β, and γ, which contrasts with catchwords internal to these sections which are usually completed by the scribe responsible for the preceding section.[18]

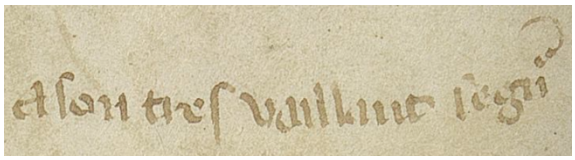
The presence of early non-scribal hands intervening in this way suggests that individuals other than the texts' scribes were responsible for the final ordering of the manuscript. The absence of final catchwords on sections ***a***, ***i***, and ***j***, meanwhile, adds to the evidence encountered throughout this chapter which suggests that hands 1, 9, and 10 are distinct from the rest of the codex. Hand 10 stands out in particular, given the unusually long text it copies (*Lucidaires*), the use of 36-line columns rather than the default 38, and the presence of quires

[18] Final catchwords do not appear in sections ***a***, ***i***, and ***j***. In section ***c***, the final catchwords by hand β appear at the end of the second to last quire, q. 25, suggesting that q. 26 was missing at the time this compilatory process was completed. As will be discussed in the section on foliation on p. 37, these later catchwords do not match any quires present in the manuscript today, suggesting the one-time presence of material which was either removed when q. 26 was rediscovered, or has been lost over the centuries. In section ***e***, meanwhile, the post-scribal addition appears alongside original catchwords in the scribal hand, reinforcing the fact that this section became fragmentary at some point after it was copied.

containing just two folios at either end of the text,[19] which all suggest hand 10's contribution was not originally contiguous with the rest of the manuscript's contents.

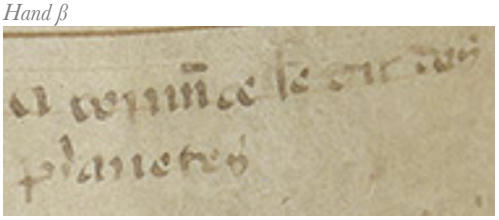
Figure 1.10 – examples of the section-final catchwords completed by the non-scribal hands α, β, *and* γ.

Hand a



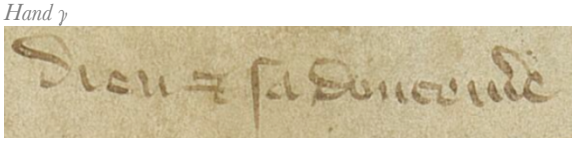
f. 56vb

Hand β



f. 262vb

Hand γ



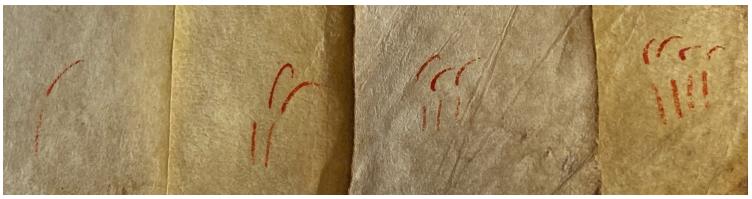
f. 246vb

Looking at the numbering of the quires within fr24432 in addition to this evidence, we find support for a different method of segmenting fr24432 which involves both the combination and the subdivision of some of the groupings given above.

There are a number of marks to be found in the codex which indicate the order that leaves and quires should appear in, many of which are apparently near-contemporary with the production of the quires themselves. Of the 59 quires in fr24432, 32 contain quire numbering marks which are seemingly in the same hand.[20] These marks are predominantly found in the bottom right hand corner of the recto side of the first 4 leaves of a quire (or 3, when a quire consists of only 6 leaves, rather than the usual 8), although many have been lost to wear and trimming. These marks consist of letters, in red ink, which proceed through the alphabet and are repeated to give a leaf's position number within the quire.[21] Thus, the first four leaves of quire 9, given the label *i*, are marked as *i*, *ii*, *iii*, and *iiii*, as in Figure 1.11.

[19] Based on the foliation, the latter quire, q. 51, appears to be missing two folios, ff. 386 and 387, since q. 51 ends on f. 385 while q. 52 begins with f. 388.
[20] Quires 2–13, 17, 19, 22, 33–42, and 52–58.
[21] We note, of course, that the medieval alphabet combines *i* and *j*, and *u* and *v*, and omits *w*.

Figure 1.11 – quire numbering marks on the first four leaves of quire 9 (ff. 63–66).



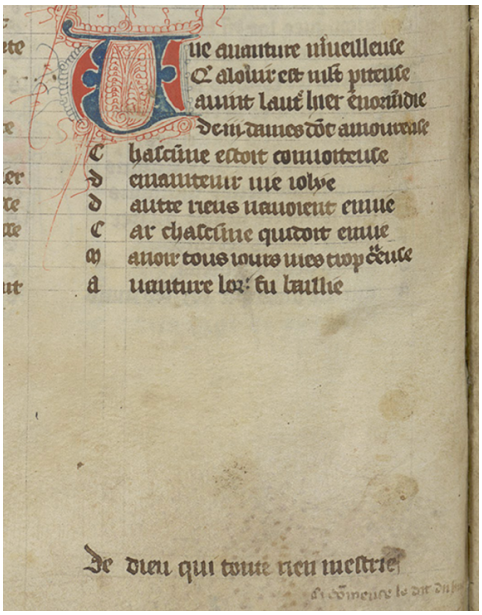
These markings run largely uninterrupted from q. 2 (labelled as *b*) to q. 22 (labelled as *y*), with those quires in the grouping for which markings are not visible (qq. 14–16, 18, and 21) nevertheless fitting into the expected sequence, since the lettering resumes in a predictable way after the expected number of intervening quires. That is, q. 13 is labelled as *n* and q. 17 is labelled as *r*, such that we can assume that qq. 14–16 were once labelled (or intended to be labelled) *o*, *p*, and *q* respectively. Given the text of q. 1 continues directly into that of q. 2, as does that of q. 22 into q. 23, we can also assume that q. 1 would have been considered quire *a*, and q. 23 may have been considered quire *z*. While qq. 24–26 contain no such markings, the fact that their texts (and scribe) are also contiguous with the preceding quires further suggests that they ought to be treated as part of the same grouping.

Later in the manuscript, beginning at q. 33 (labelled as ~~*b*~~), we find another set of quire numbering marks in the same hand and ink, this time struck through, as in the example of ~~*i*~~ on the fourth leaf of q. 40, given in Figure 1.12. Given that the grouping begins with a quire labelled ~~*b*~~, we could assume that the preceding quire, q. 32, has simply been left unmarked (as with q. 1), but given the scribal structure it in fact appears more likely that it was originally q. 30 that preceded q. 33, with qq. 31 and 32 constituting a later insertion. This allows us to identify a continuous stint by hand 7 beginning in q. 27, immediately after the end of our first set of quire marks, suggesting that the quire markings with and without strikethrough marks once ran contiguously from one to the next.

Figure 1.12 – quire numbering marks on the fourth leaf of quire 40 (f. 306).

Indeed, if qq. 31 and 32 are out of their original place, then this helps to explain the fact that the final item of q. 32 (*Trois mortes*) is fragmentary, as well as the presence of two catchwords, visible in Figure 1.13. One of the two catchwords, in the same hand as the text (hand 3), refers to the next line of the poem, and therefore to its original following quire, while a non-scribal but roughly contemporary hand (α) corrects the original catchwords with the first line of q. 33, to provide continuity in fr24432. The two quires now forming qq. 31 and 32 were therefore likely taken from elsewhere (perhaps even from within another manuscript) and were included at this position in fr24432 on the basis of their content, with the quire following q. 32 in its original context not considered worthy of inclusion in this position, thus resulting in the fragmentation of *Trois mortes*.

Figure 1.13 – the fragmentary text Trois mortes *(f. 246ᵛᵇ).*



In the case of this second, struck-through set of quire markings, we note that unlike the first set of markings, the run of marked quires is not continuous. The letters run as expected from q. 33 (~~*b*~~) to q. 42 (~~*l*~~), before disappearing on qq. 43–51, and then reappearing from ~~*n*~~ on q. 52, and running uninterrupted to the final quire of the manuscript (q. 58, ~~*t*~~). The presence of an intervening run of 9 quires is therefore significant, and can give us a clear indication that these quires were added later: after this quire marking was complete, an original quire labelled ~~*m*~~ was removed, and the later quires were added.

It is this that explains the brevity of q. 43, at only 2 singleton leaves, since when quire ~~*m*~~ was removed, and with it the end of the text begun in q. 41 which runs through q. 42 (*Messe*), the end of the text was recopied by hand 9 to allow the text to be completed, with an intervening blank folio then inserted before the beginning of another contiguous grouping of quires, qq. 44–51, which only contain one text (*Lucidaires*) in the work of a single scribe (hand 10). Since the remaining text of *Messe* runs to only one leaf, if we assume that the quire labelled ~~*m*~~ ran to the standard length of 8 leaves, then there must once have been 7 leaves of content that were discarded, and are now lost to us.

Table 1.10 – groupings of quires in fr24432 by quire marking.

Grouping	Quire(s)	Scribal hand(s)	Item(s)
A	0	1	Table of contents
B	1–26	2, 3, 4, 5, 6	1–52
C	27–30	7	53–56
D	31–32	8, 3	57–60
E	33–42	7, 5	61–77
F	43	9	77
G	44–51	10	78
H	52–58	2, 11	79–90

If we follow this method of segmentation, and combine this with the observations made on the basis of the discrete textual content in Table 1.9, we are therefore able to update our division of fr24432 into at least seven primary groupings of quires, as in Table 1.10, with

groupings ***E*** and ***H*** originally forming part of the same contiguous unit, and ***F*** and ***G*** constituting connected but distinct groupings which were added later.

The presence of hand 3 in grouping ***D***, hand 5 in grouping ***E***, and hand 2 in grouping ***H*** gives sufficient overlap to mean that hands 2–8 and 11 worked together closely, suggesting that they worked as part of a single, large *scriptorium*. The presence of additional catchwords hands in between and within these groupings suggests an even broader set of individuals involved in the manuscript's production than just these 8 scribes, as well as the fact that the groupings identified in Table 1.9 were probably those according to which the quires were copied, before they were quickly assembled into the groupings identified in Table 1.10.

Hands 1, 9, and 10, meanwhile, do not overlap with the surrounding scribes, such that we may be tempted to assume that the work of hands 9 and 10 in particular was added to fr24432 once the manuscript had left its *scriptorium*. In fact as we have already seen for hand 9, and as we shall soon see for all three scribes, the inclusion of their work into fr24432 appears to have been finished before the manuscript was deemed complete and ready for sale or delivery to a patron, thus suggesting that all 11 scribes were present in the same *scriptorium*, albeit not necessarily at the same time.

Comparing these revised groupings to those we gave initially in Table 1.9, we note that the subgroupings ***b*** and ***c***, as well as ***f*** to ***i***, and ***k*** and ***l*** are combined into broader groupings. This is not to deny the fact that these subgroupings had some initially distinct material reality, as evidenced in particular by the presence of the non-scribal catchwords hands. They were clearly brought together by someone other than the scribe(s) who copied them. The reason for combining these subgroupings is simply that this material distinctness appears to have been very short lived, with the quires combined soon after their production in a way that indicates that they were originally planned in this order. Our attention is therefore better

focussed on identifying those groupings which appear not to have been originally conceived as a single unit.

1.1.4 Pricking and ruling

Further evidence in support of the division of fr24432 into distinct groupings with somewhat diverse origins can be found in the way that the manuscript's leaves are pricked and ruled in preparation for the addition of text. Since the manuscript has been trimmed, pricking marks are not always visible, but there are enough surviving marks to allow us to draw some broad conclusions.

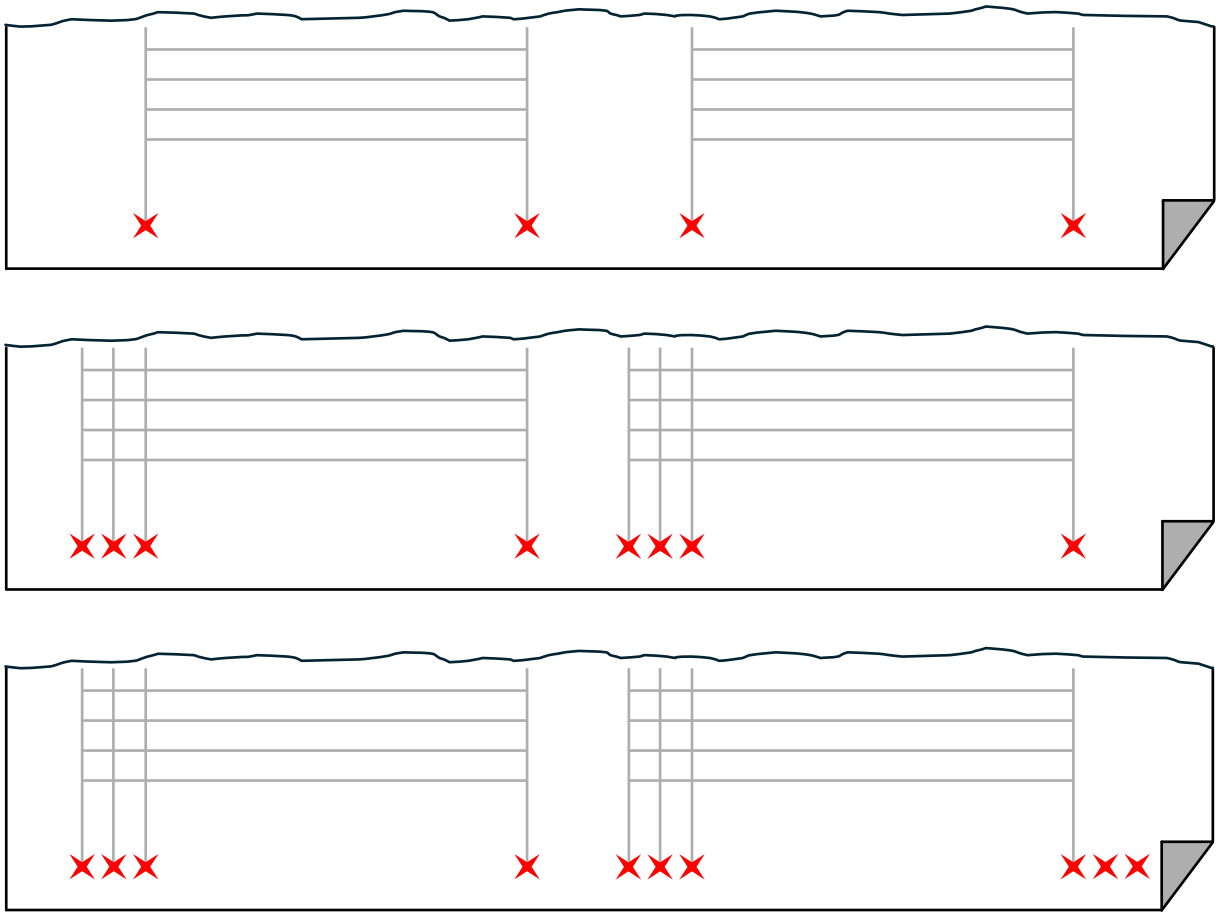
The pricking in the codex has been completed using one of two methods: either by piercing the parchment with a pin, or by using a knife to create a small slit, which may be vertical, horizontal, or diagonal. The number of vertical pricking marks is of course determined by the number of lines required for a given leaf,[22] which we have considered in detail above, but there is also notable variability in the number of horizontal pricking marks. These are used to determine the position of the vertical lines which define the limits of the text block and its internal division into columns, and in fr24432 they appear in combinations of 4, 8, and 10 marks.

When 4 horizontal pricking marks are used, they simply set out the left and right extremities of each column. In the sets of 8 marks, there are 3 at the beginning of each recto column to mark the space by which to separate the initial grapheme from the rest of the line, and a further mark to indicate the right edge of each column. Sets of 10 marks, meanwhile, overcome a problem with the 8-mark system whereby the pricking is less useful on the verso than on the recto, by adding a further two marks on the far right hand side of the recto, and therefore the far left hand side of the verso. These therefore allow the same vertical lines to be

[22] When the vertical marks are made with a pin, the bottom mark is doubled.

drawn on the verso as on the recto to permit separation of the initial grapheme of a line from the remainder of the text. These three pricking methods are illustrated in Figure 1.14.

Figure 1.14 – pricking layouts found in fr24432 of 4, 8, and 10 marks, as seen from the recto side.



The distribution of these different horizontal pricking layouts cannot be straightforwardly linked to the groupings we have identified above, but two general principles hold true, with limited exceptions:

1. The method and layout used for pricking is consistent within the same quire, indicating that whole quires were pricked and ruled at once, before text was added.

2. Scribes have a preference for the method and layout that they use, indicating some scribal involvement in the pricking and ruling process.

Thus, when they are the primary scribe in a given quire, hands 1, 3, and 8 consistently make use of slit marks created with a knife.[23] Hands 5, 6, and 7, meanwhile, are consistent in their preference for the use of a pin, and hands 2 and 11 are more variable, sometimes using a slit, and other times a pin hole. No pricking marks survive in the work of hands 9 and 10.

Horizontal pricking marks like those above only appear in the work of hands 2, 5, 6, 7, 8, and 11. Hands 2, 6, 8, and 11 consistently make use of the 4-mark approach, and hand 5 sometimes also uses the 8-mark approach. Hand 7 shows examples of all three approaches, with 8 marks appearing on the quires of their first stint (qq. 27–30), and then their second stint beginning with sets of 10 marks (qq. 33–38) before they turn to sets of 4 marks for q. 39 (the marks have been lost from q. 40, the final quire of the stint).

Figure 1.15 – groupings of quires in fr24432 with details of pricking.

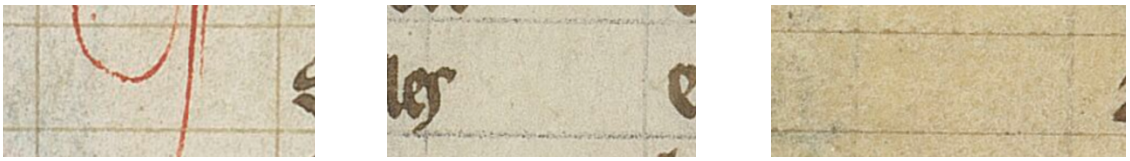
Grouping	Quire(s)	Scribal hand(s)	Pricking Type	Pricking Layout
A	0	1	Slit	–
B	1–26	2, 3, 4, 5, 6	Pin / Slit	8 / 4
C	27–30	7	Pin	8
D	31–32	8, 3	Slit	4
E	33–42	7, 5	Pin	10 / 8 / 4
F	43	9	None	
G	44–51	10	None	
H	52–58	2, 11	Slit / Pin	4

While there is some variation within the groupings that we identified above, there are nevertheless general trends which can be identified that support our overarching hypothesis about fr24432's internal diversity. Groupings ***C***, ***D***, and ***E*** are all internally consistent in their methods of pricking, and for the most part also in their horizontal pricking layouts, but grouping ***D*** in particular appears distinct from those surrounding it, reinforcing the suggestion that this grouping constitutes a later insertion. That groupings ***F*** and ***G*** both

[23] Hand 3 appears within a pin-marked quire for which hand 2 was the principal scribe, suggesting that hand 3 was not responsible for putting the quire together. We must ignore hand 4 since they are not the primary scribe in the quire in which they appear (q. 10).

contain no evidence of pricking marks, while possibly a simple accident of the process of trimming, also reinforces our sense that these two groupings were added to fr24432 at the same time, at a moment distinct from the compilation of much of the rest of the codex.

Figure 1.16 – from left to right: wax crayon ruling (f. 16ᵛ); lead ruling (f. 79ʳ); mixed lead and ink ruling (f. 208ʳ).



A similar story is told by the materials used in the ruling of the quires of fr24432. Quires 0–8, 31, and 43–55 are ruled using wax crayon, while lead ruling appears in quires 9–27, 32, 35–42, and 56–58.[24] Quires 28–30 and 33–34, meanwhile, are ruled with lead vertical lines, and ink horizontal lines.

Figure 1.17 – groupings of quires in fr24432 with details of ruling.

Grouping	Quire(s)	Scribal hand(s)	Ruling Material
A	0	1	Wax
B	1–26	2, 3, 4, 5, 6	Wax / lead
C	27–30	7	Lead / lead-ink
D	31–32	8, 3	Wax / lead
E	33–42	7, 5	Lead / lead-ink
F	43	9	Wax
G	44–51	10	Wax
H	52–58	2, 11	Wax / lead

There is some degree of variation by scribe at play here, for example hand 2 consistently writes on quires ruled with wax crayon, while hand 6 only appears on leaves ruled in lead. The clearest picture comes however when we take a somewhat more macro-level approach, by mapping the use of these materials onto our groupings of quires. When we do so, as represented in Figure 1.17, a clear picture of their use emerges that fits with our overarching hypothesis about the diversity present in fr24432.

[24] The ruling on q. 41 has faded to the point that it is no longer visible, but given the surrounding quires (some of which also contain particularly faint ruling) it is likely that it was once ruled in lead.

Groupings ***B***, ***D***, and ***H***, containing the work of hands 2, 3, 4, 5, 6, 8, and 11, appear to be unified in that they contain a combination of the use of wax and lead ruling, thus reinforcing our impression that the quires of these groupings were likely produced in the same location, with the diversity of ruling material preferred by individual scribes nevertheless forming a coherent whole. Most striking, perhaps, is the fact that groupings ***A***, ***F***, and ***G*** make consistent use of wax ruling, contributing to our sense of their collective difference from the remainder of the codex, and reinforcing the sense that they constitute a later addition.

In groupings ***C*** and ***E***, meanwhile, which centre around the work of hand 7, we find quires ruled in lead and the lead-ink pairing, again reinforcing the sense that these groupings are somehow connected, and providing further evidence for our hypothesis that grouping ***D*** was moved to this position later. In this case, from the overlap between the ruling materials used in grouping ***D*** and those found in groupings ***B*** and ***H***, and of course the fact that hand 3 appears in groupings ***A*** and ***D***, we can surmise that if grouping ***D*** was indeed removed from another codex to be inserted in fr24432, this codex was nevertheless the work of the same *scriptorium*.

On a more general level, from the internal consistency of individual quires, and the clear association between scribal identity and pricking and ruling methods, we can state with some certainty that the quires of fr24432 were fully pricked and ruled before the text began to be added, and that the preferences of the scribe who copied the text had an impact on the preparation of the text block. de Hamel (2018: 58) notes that 'by the fourteenth century it seems to have been possible to purchase gatherings of parchment ready prepared for writing,' and so it is possible that the scribes did not lay out the text block themselves, rather they chose pre-fabricated quires that suited their individual requirements.

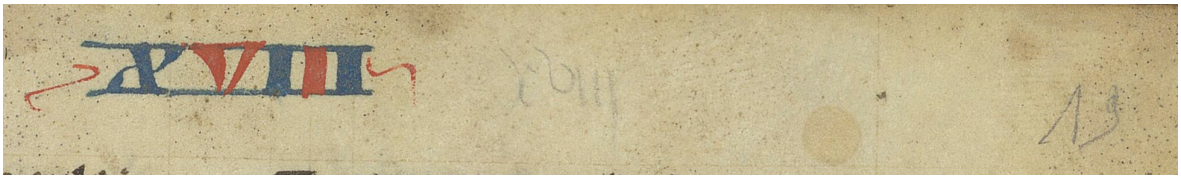
What remains most notable of all is simply the range of approaches and materials used in the same codex. A number of different methods and layouts of pricking and ruling were used interchangeably within the same *scriptorium* or their parchment suppliers, revealing an acceptance of material diversity amongst those behind the production of fr24432.

1.1.5 Foliation

Thus far, we have seen the extent to which fr24432 is an internally diverse manuscript. It is not simply the case, however, that it is a composite collection of different manuscripts produced at different times, and at different *scriptoria*, since there is still a great deal of evidence of unity that appears to date from very early in the manuscript's existence. As a result, we are able to nuance our understanding of the diversity identified above by concluding that it already existed before fr24432 left its place of production.

Our first example of this apparent unity comes from the manuscript's contemporary foliation. Fr24432 contains two main rounds of foliation, consisting of modern pencil markings alongside much earlier numbering which appears to be nearly contemporary with much of the manuscript's textual content. This earlier foliation consists of large Roman numerals in a single hand at the head of the recto of folios in alternating red and blue ink, with the notes from which the foliator worked sometimes still visible.[25]

Figure 1.18 –early foliation in Roman numerals, the notes used by the foliator, and modern Arabic numeral foliation (f. 18r).



[25] There are occasional errors in the foliation, particularly between the Roman numerals *L* and *I*, which indicates that the scribe responsible for the foliation was likely not the person who wrote the guiding notes.

Some folios have seemingly been lost after the original foliation was completed. For example, the loss of f. 48 from the end of q. 6 is indicated both by a jump in the textual content and in the foliation. This thus places the foliation as having been completed relatively early in the manuscript's life.

There are also reasons to suspect, however, that this foliation is not exactly contemporary with when the body of the codex was copied, with at least some degree of modification within the codex having occurred before it was foliated and released from the *scriptorium*. One quire (q. 26), labelled in the modern foliation as ff. 198A–G, has seemingly been missed by the early foliator, since no pagination exists in this hand, with the previous quire (q. 25) ending on f. 198 and the following quire (q. 27) beginning with a folio numbered as f. 199, despite q. 26 representing an otherwise unremarkable continuation of both the text (*Tournoiement*) and hand (6).

Figure 1.19 – catchwords at the end of quire 25 (f. 198^{vb}).

Notably, at the end of q. 25, the catchwords in the scribal hand, which have been damaged by trimming but apparently match the first line of q. 26 on f. 198A^{ra}, are accompanied by a non-scribal hand in a lighter ink, seemingly roughly comparable in terms of date, which proposes 'pour ceuls' as the phrase which will begin the following quire.[26] This phrase does not appear verbatim anywhere in the manuscript, and no quire begins with a phrase that

[26] It is possible that this hand is the same as the hand α identified on p. 27 above.

could be seen to match this suggestion,[27] perhaps indicating the temporary loss of q. 26, combined with the temporary erroneous inclusion of an alternative quire which was later removed from the manuscript. Over the period during which the original foliation was completed, however, neither quire appears to have been in place, given the continuous foliation between q. 25 and q. 27, which points to at least two phases of change occurring to this section of fr24432 in its early life, both likely before it left the *scriptorium*.

Further evidence that there were processes of recompilation which occurred before the foliation comes from grouping ***D***, consisting of quires 31–32. As we have discussed, the quiring and scribal structures appear to suggest that these quires were inserted after the original compilation of the codex was complete, with the inclusion of these quires resulting in the fragmentation of *Trois mortes*, the text that appears at the end of q. 32. The foliation here carries on regardless, with the foliator responsible seemingly unaware, or at least unconcerned, that there is any gap within the material at all. Similarly, the foliator also continues the pagination uninterrupted into groupings ***F*** and ***G***, despite the fact that, as we have seen, these groupings are apparently distinct from the rest of the codex and constitute a later insertion.[28]

The necessary conclusion of this is that there were processes of recompilation or loss which occurred shortly after much of the textual content of fr24432 had been copied, and all of which took place before the manuscript left its *scriptorium*.

---

[27] There are four analogous phrases, none of which have comparable spelling, and none of which appear at the beginning of a quire: 'pour ceulz' (f. 152ra), 'pour ceus' (f. 331ra) (appears twice), and 'pour ceulz' (f. 372^{vb}).
[28] The fact that two folios are foliated as 328, one at the end of grouping ***F*** and one at the beginning of grouping ***G***, appears simply to be an error on the part of the foliator, who is no stranger to mistakes elsewhere in the codex, often involving the omission of individual numerals. The jump in the foliation between f. 385 and f. 388, as we have seen, results from the removal of two blank folios at the end of grouping ***G***, and therefore likely dates later than the period of the manuscript's production.

1.1.6 *Litterae notabiliores*

The 4,203 *litterae notabiliores* found in fr24432 provide further indications that the processes of recompilation affecting the codex took place before its completion and release from the *scriptorium*. The *litterae notabiliores* vary in size and prominence, from inline pilcrows to initials spanning six orthographic lines, suggesting some internal diversity, but they do nevertheless appear to be the work of a single flourisher.[29] The smallest of these, in particular the pilcrow and initials of a height of one or two lines, are used primarily for subdivisions within longer sections, whereas taller initials, especially those over two lines in height, are generally used to mark more significant shifts within the manuscript, such as separations between texts.

Figure 1.20 – representative examples of the litterae notabiliores *found within fr24432.*

Pilcrow

f. 398rb

Single-line

f. 237rb

Two-line

f. 256vb

Three-line

f. 152vb

Four-line

f. 19ra

Five-line

f. 17ra

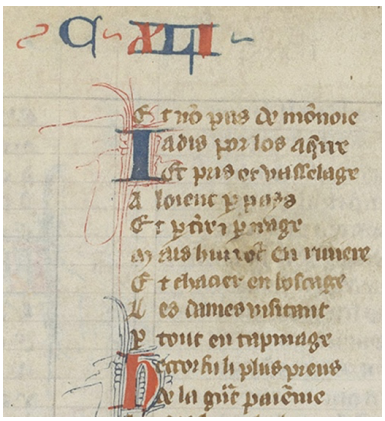
Six-line

f. 23ra

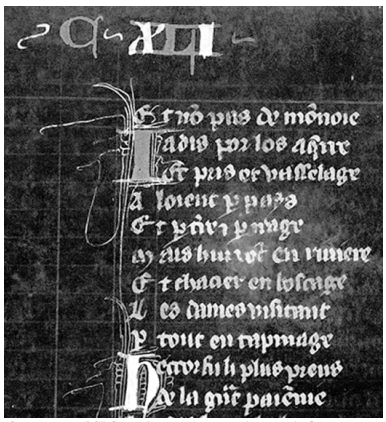
[29] The flourisher is a close (but not exact) match for Rouse & Rouse's Flourisher E (2000, i: 66), who appears in BnF lat. 395, the fourth volume of a multi-volume Bible produced in Paris by Gui de Clermont and his *atelier*.

In most cases, bar a few apparent inconsistencies, the colours used for the initials alternate between red and blue, with the alternative colour used for extra decoration in those *litterae notabiliores* with a height of two or more orthographic lines. The two inks used, as well as the general style of the characters, appear to be largely comparable with the foliation that appears throughout the manuscript,[30] suggesting that these features were added by the same *scriptorium*, and perhaps even the same individual.

Figure 1.21 – representative example of foliation and litterae notabiliores*, with spectral analysis using* retroReveal *(Erickson* et al.*, 2011).*



f. 141rb



f. 141rb (YIQ colour space, channel Q)

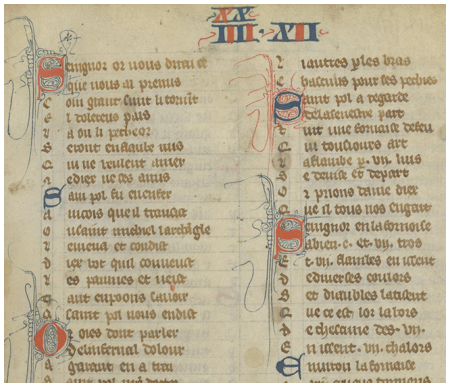
That said, the foliator and the flourisher, even if they were the same person, did not work on the manuscript at exactly the same time, given that in quire 26 (f. 198A–198G), which was missed by the foliator, the *litterae notabiliores* continue uninterrupted, with no indication of a change in hand. This implies some time gap between the inclusion of the *litterae notabiliores*, which perhaps came first, and the foliation, which perhaps came after the loss of the quire.

The frequency of the use of *litterae notabiliores* appears to be determined more by the form of a text than by any other factor, with strophic verse often using one initial per stanza, while non-strophic verse and prose use considerably fewer. Variation in height, meanwhile, is more complicated, and in some cases is somewhat correlated with scribal preference. There are two

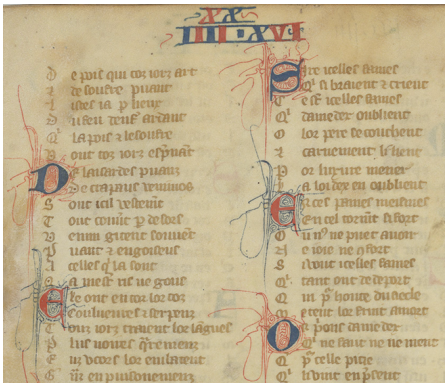
[30] Based on image analysis using the online tool *retroReveal* (Erickson *et al.*, 2011), there appear to be no visible differences between the inks used in the foliation and those used for the *litterae notabiliores*.

cases in which the preferences of the scribe copying a given section of text appears to lead to variation between the heights of the initials, both of which involve hand 3. For *Vision* in q. 12, hand 3 uses a combination of single-line and two-line initials, but when hand 5 takes over the same item at the beginning of q. 13, this variation ends, and two-line initials are preferred as the default method of subdividing longer sections of text.

Figure 1.22 – representative example of litterae notabiliores *in quires 12 and 13.*



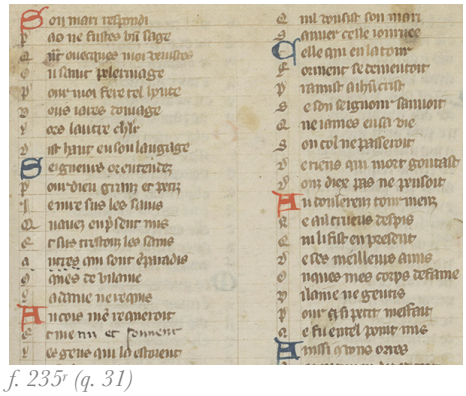
f. 92r (q. 12)



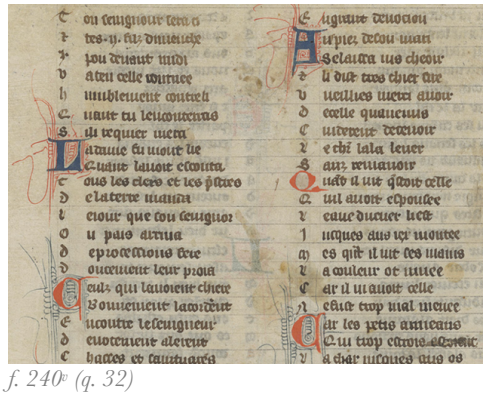
f. 96r (q. 13)

Similarly, when hand 3 takes over *Anelés* from hand 8 at the beginning of q. 32, the hitherto consistent use of a single-line initial is muddied, with a two-line option preferred but accompanied by instances of single-line initials. Both these examples therefore indicate that for certain texts hand 3 was comfortable with variable usage of single- and two-line initials, or perhaps that they simply paid less attention to the gaps they left for the flourisher to fill, whereas other scribes were more concerned with consistency.

Figure 1.23 – representative example of litterae notabiliores in quires 31 and 32.



f. 235r (q. 31)



f. 240v (q. 32)

It is also possible that this variation was present in the exemplar from which the scribes copied these particular texts, and that scribes other than hand 3 saw fit to regularise it, whereas hand 3 did not, which would explain the consistency seen in the initial heights of hand 3 in other sections of the codex, should the exemplar(s) for those texts have been more consistent. If so, this would raise the possibility that *Vision* and *Anelés* were copied from the same exemplar, interesting given that there is little to directly connect them, other than the fact that they are both written in rhyming alexandrine quatrains.

Besides these examples, most variation in initial height does not appear to be determined by the scribal hand present in that section. A clear example of this is to be found in the long stint of hand 6, which sits within grouping ***B*** and covers 13 quires (14–26) and 24 items, of which the first (*Vie du monde*) is a continuation of a text begun by hand 5, and the last is fragmentary (*Lai d'amour 1*). Throughout this stint, the most common initial used to separate subsections of longer text has a height of two lines, however there are examples when variation occurs for individual texts. *Chevalier et hermite* and *Cordouanier*, for example, vary between single-line and two-line initials with a preference for the latter, whereas *Isopet* displays variation with a preference for the former.

This variation within scribal stints could help us isolate certain texts which could have come from the same exemplar manuscript. Across the codex, we see that a two-line initial is by far the most frequent means of indicating the subdivision of longer text, with 56 texts preferring this method outright. Besides the table of contents, single-line initials are preferred in *Vin*, *Pommes*, *Espervier*, *Merlin*, *Robert*, and *Jouenne*, dominant with some variation in *Bible*, *Isopet*, and *Division*, and dominant or preferred in one of the two hands behind items *Vision* and *Anelés*, as discussed above.

From this, two clusters of texts emerge in particular, namely *Pommes*, *Espervier*, *Bible*, and *Vision*, which appear in positions 20, 23, 24, and 26 within the codex, between ff. 53va and 99ra, as well as *Isopet*, *Merlin*, *Robert*, *Anelés*, and *Divisions*, which appear as items 50, 53, 54, 57, and 59, from ff. 171^{rb}–246^{vb}. The presence of intervening texts with different tendencies between the members of the former group can be explained by the fact that *Aristote*, appearing at item 21, is a prose text with different requirements for initials to the surrounding verse texts, while *Planetes 1* and *Dormant* (items 22 and 25) contain no subdividing initials. It is therefore possible that this whole run of texts was copied from a single exemplar manuscript which was itself distinct from those which were the exemplars for the texts surrounding this section.

Examining those texts within the latter group which do not prefer single-line initials, we find that *Lai d'amour 1* (item 52) is fragmentary and so offers little evidence of its preferences, and *Yonec* (item 58) is non-strophic verse, unlike many of the other texts in this section, and therefore has too few *litterae notabiliores* to discern a clear preference. That said, in this case not all of the exceptions can be explained away so simply. *Tournoiement*, *Florence*, and *Menage* (items 51, 55, and 56) are consistent in their use of two-line initials, in contrast to the texts that surround them, suggesting this cluster of texts may have been copied from a number of different manuscripts.

Robert and *Florence* (items 54 and 55) are particularly notable, since both texts have comparable form (alexandrine quatrains) and content (rewritings of longer epic tales), and appear in the same scribal stint (hand 7) and even the same quire (q. 29). They even, as we shall argue in Chapter 3, are the work of the same author. And yet, both texts differ in their preferred initial. It is therefore arguably possible that *Florence* was copied after *Robert* exactly because of these similarities in form and content, but that it was copied from a different exemplar manuscript with different preferences on the height of initials. If true, this would

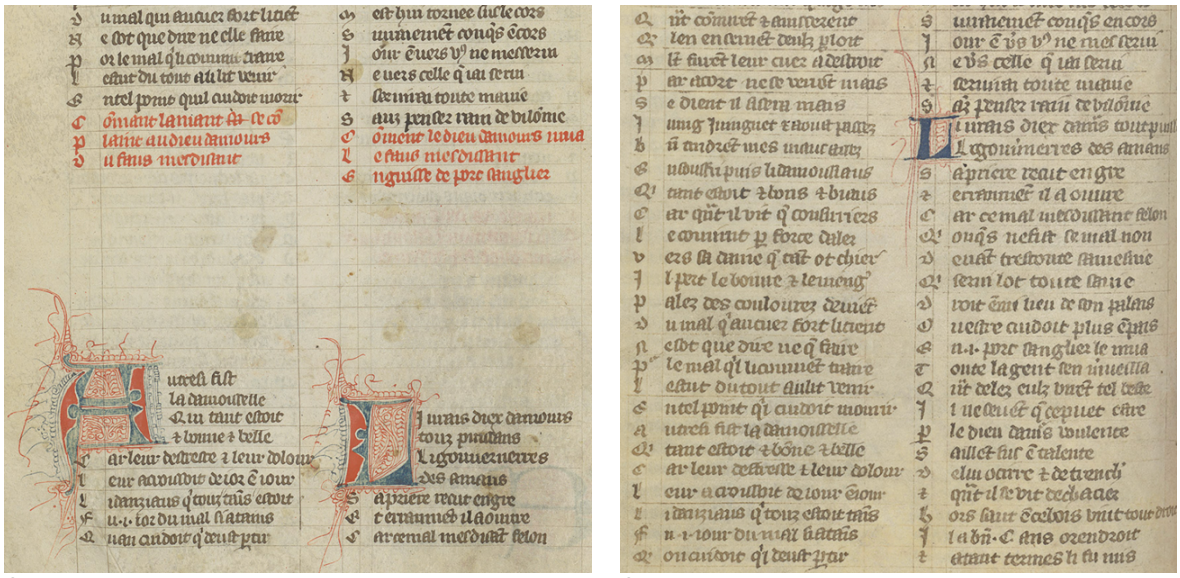
reveal the importance of both content and form as organising principles in the manuscript's compilation.

We must of course entertain the possibility that this variation in initial height was already present in a single exemplar from which the scribes copied. If this were the case, then we would be able to make a similar point here but for a manuscript further up the exemplar chain than the present codex, namely that a decision was made at some point to put these two texts alongside each other, perhaps due to their similar form and content, despite originally being held in different manuscripts.

Alongside conclusions that can be drawn from those texts that prefer single-line initials, the limited use of the pilcrow within the manuscript is also revealing. The pilcrow is most common in *Songes* and appears occasionally within *Salomon*, in addition to very sparse use elsewhere. Both these texts are in prose, and are written in a list-like way that suits frequent divisions, but neither feature is unique to these texts, and therefore for *Songes* in particular, this suggests that the text was copied from an exemplar manuscript which was not used for most other texts within the codex.

In *Mesdisans 1*, a non-strophic text in octosyllabic verse appearing at item 4, we observe the unusual use of a four-line initial to distinguish between relatively minor subsections of text, whereas it is seen elsewhere usually as a means of separating texts or major sections within texts. Interestingly, the same is not true of the second copy of the same text found at item 86 (*Mesdisans 2*), which instead prefers the much more common two-line initial, if it separates subsections at all, despite seemingly being from the same exemplar given the minimal textual differences between both copies of the text.

Figure 1.24 – the same section of text in both copies of La chace des mesdisans.



f. 18v *f. 408v*

A possible explanation for this could be seen in the fact that *Mesdisans 1* also contains spaces for illustrations which were never completed, along with rubricated captions, and as such it is possible that the exemplar from which the scribe behind both copies (hand 2) was working was much more lavish than the exemplars of many other texts in the codex. This may have been imitated by hand 2 in the first copy, perhaps in the expectation that the present codex would be as lavishly illuminated as the exemplar, but the second copy was produced with lesser expectations as to the manuscript's level of decoration.

The above analysis is therefore revealing of the possible exemplars consulted by the scribes of fr24432 when copying its texts. For our overarching argument, however, the most important observation regarding the *litterae notabiliores* remains their visual consistency across the codex. While there is variation in the size and frequency of the forms, the fact that the spaces left for embellishment are determined by the scribe means that this variation results from scribal preference, possibly under the influence of the exemplars used and even the nature of the text being copied, rather than due to a variation in the identity of the flourisher. We can therefore

say with some certainty that the flourishing was some of the last work to be completed within fr24432, and was all complete before the manuscript left the *scriptorium*.

1.1.7 The table of contents

One final object of consideration suggesting that at least some unity was achieved within fr24432 before it left its *scriptorium* is the table of contents. This occupies three columns of f. 0, the only leaf of q. 0, with the fourth column left blank. The parchment is in poor condition, consistent with having been at the front of the manuscript for much of the last seven centuries. It is largely accurate in its titles and page references, corresponding in most cases to the earliest foliation, although some titles do differ from those in the rubrication surrounding the texts,[31] and *Dormant*, *Lai d'amour 1*, *Trois mortes*, and *Mesdisans 2* are missed out. The omission of item *Dormant*, a very short text written entirely in red, is unsurprising since it may not have been considered an independent text even if it was present when the table of contents was created (which is likely, given it is in the same hand as the surrounding text).

The omission of *Lai d'amour 1* and *Trois mortes* is more interesting, and can tell us a good deal about when the addition of the table of contents took place within the timeline of the manuscript's production, since it enables us to better understand when the ends of these texts were lost. As we have seen, *Lai d'amour 1* is fragmentary as a result of the loss of the final folio of q. 26, but the omission of this text from the table of contents more likely stems from the fact that the entirety of q. 26 was lost for a period within the production of the codex, during which point the foliation was also completed. This therefore places the addition of the table of

[31] One interesting example is *Boivin de Provins*, which is the only text contained within the manuscript generally recognised to be a *fabliau*. The male trickster's name is used as the title in the scribal *explicit* of the text in the body of the manuscript, as indeed is the case in the other manuscript in which the text appears, BnF fr. 837 (*fabliau* manuscript A). In the table of contents, however, the title is given as *De mabille de provins*, with the text referring to the persona adopted by the text's female trickster. This is of note since the male trickster is the only person explicitly said in the text to be from *Provins*, and the name Mabille does not appear towards either end of the text, and is not highlighted in any way when it does appear. The scribe responsible for the table of contents therefore appears to have read the full text to look for an internal title, given the absence of an incipit, and settled on one of their own invention which they deemed appropriate.

contents as after the completion of the foliation, but before the addition of the *litterae notabiliores*, which was achieved after q. 26 was returned, lacking its final folio.

Further, in omitting *Trois mortes* from the table of contents, the scribe reinforces our sense that it was a process of recompilation involving the insertion of grouping ***D*** that led to this text becoming fragmentary, rather than a post-*scriptorium* removal of material from fr24432. The scribe consciously chose to overlook this text, whose partial inclusion in the codex was clearly considered an accident of the recompilation process rather than an intentional decision, thus indicating some awareness of the compilatory decisions behind this reorganisation of quires.

Of similar interest is the omission of a reference to *Mesdisans 2*, which is a duplicate of the much earlier *Mesdisans 1*. It is possible that this is a recognition, or perhaps an assumption, on the part of the scribe, that the duplication of this text was a mistake. It is therefore interesting to note that the other text which appears twice in the manuscript, *Planetes*, does have two references in the table of contents, so perhaps our scribe simply did not notice the duplication of this text, or it was considered intentional.

One entry in the table of contents appears to have been erroneously included. It refers to a text entitled *Des .ii. amans* appearing on f. 16 (f. 0^{r}), which does not correspond to a separately divided text, or indeed to any relevant subsection of the text found on that folio, *Respon*. An early modern hand has struck through the entry and asserted that it '[n']existe pas', but a possible explanation does reveal itself from f. 18ra, in which a rubricated heading states:

> Ci parole des deus ama*n*s
> Qui s'entramoie*n*t si loiaume*n*t
> Q*ue* mesdisa*n*t fist dessevrer.

This heading, and the text that follows it, is simply a subsection within *Mesdisans 1*, and so the scribe responsible for the table of contents appears to have misinterpreted the rubricated

heading as an incipit, since it is visually similar to that found at the beginning of *Mesdisans 1* on the previous folio. The inconsistency in the folio reference therefore simply results from a misreading of the scribe's notes.

Similarly, in the case of the verse allegory *Arbre*, the table of contents sections the text differently to the body of the manuscript. In the table of contents, the text is divided in three, with *De l'arbre d'amours* said to begin on f. 265, which matches the contemporary foliation, and *La branche de biauté* and *La branche d'orgueil* said to begin on folios 267 and 268 respectively. The first of these subdivisions makes some sense, since it corresponds to a gap of 12 orthographic lines within the text on f. 267^{vb}, which appears originally intended to contain an illustration of the text's contents, given the presence of a rubricated caption, and so it is possible that as with *Mesdisans 1*, the scribe erroneously took this to be the end of one text and the start of another.

The inclusion of a second subdivision on f. 268, however, appears to be a less logical error, since in this case the reference to *La branche d'orgueil* appears only in a two-line rubricated heading on f. 268ra, rather than being separated from the surrounding text by any empty orthographic lines. Perhaps the most likely explanation is therefore that the scribe was influenced by the perceived textual division on the preceding page, and took the rubrication to be another such separation. The absence of any further divisions of this text, despite a number of other such rubricated headings, suggests that upon turning to f. 268^{v}, the scribe discovered two further headings with no physical separation, and decided against including them all in the table of contents.

These apparent errors are important in that they suggest that the scribe responsible for the table of contents completed their work by referring directly to the manuscript, rather than to any draft document finalising the manuscript's order. It also points to the presence of a time gap between the creation of the bulk of the manuscript's contents and the table of contents.

1.2 Production and recompilation

Where does the seemingly contradictory evidence that we have seen thus far, of both internal diversity and unity, leave us regarding the sequence in which fr24432 was produced? If we combine all the observations made to this point, we reach a clear picture of the manuscript's production. The quire groupings are repeated here in Table 1.11 for ease of reference.

Table 1.11 – groupings of quires in fr24432.

Grouping	Quire(s)	Folio(s)	Scribal hand(s)
A	0	0	1
B	1–26	1–198G	2, 3, 4, 5, 6
C	27–30	199–230	7
D	31–32	231–246	8, 3
E	33–42	247–326	7, 5
F	43	327–328	9
G	44–51	328(bis)–385	10
H	52–58	388–443	2, 11

Based on the material, scribal, and textual evidence, it is possible to identify 12 distinct phases in the production of fr24432 before it left its *scriptorium*, namely the following:

1. The base manuscript (groupings ***B***, ***C***, ***E***, and ***H***) is copied by scribes 2, 3, 4, 5, 6, 7, and 11.

2. Grouping ***D***, containing the work of hands 8 and 3 is inserted between the previously contiguous grouping ***C*** / ***E***, having been removed from a different context, most likely a codex other than the base manuscript of fr24432, which was nevertheless produced by the same *scriptorium*.

3. A quire labelled ~~*m*~~ is removed from the previously contiguous grouping ***E*** / ***H***.

4. Grouping ***F*** is copied by hand 9 from the removed quire labelled ~~*m*~~ to complete the text with which the previous quire ends (*Messe*).

5. Grouping ***G***, containing just *Lucidaires*, is copied by hand 10.

6. Groupings ***F*** and ***G*** are inserted in the previously contiguous grouping ***E / H***, in place of the removed quire labelled ~~*m*~~.

7. Quires are removed from grouping ***H*** at the end of the codex, perhaps to make space in the binding for the newly added groupings ***F*** and ***G***.

8. Quire 26 is temporarily misplaced from the end of grouping ***B*** and is replaced by another quire beginning 'pour ceuls', which is removed shortly after erroneous catchwords are added.

9. Foliation is added to the full codex.

10. The table of contents is compiled by hand 1 and added to the beginning of the codex as grouping ***A***.

11. Quire 26 is replaced in its original position at the end of grouping ***B***, having lost its final folio.

12. The *litterae notabiliores* are added to the full codex.

Some of the ordering of these events is tentative. For example, the removal of material from the end of grouping ***H*** could have occurred before the insertion of groupings ***F*** and ***G***, although the fact that the table of contents ends with the manuscript's final item suggests that it is unlikely to be a much later modification. Similarly, it is possible that the insertion of groupings ***F*** and ***G*** predates the insertion of grouping ***D***, but given the continuity of hands between grouping ***D*** and the base manuscript, this does appear unlikely.

We may choose to ask why such processes of compilation were considered worthwhile. In the case of grouping ***D***, this could be on account of the presence of *Anelés* which fills all of q. 31

and extends into q. 32, since the text shares a great many formal, textual, and thematic similarities with *Merlin*, *Robert*, and *Florence*, which appear earlier in grouping ***C***.[32]

As for the inclusion of groupings ***F*** and ***G***, enabling the addition of *Lucidaires*, grouping ***G*** was produced as a distinct booklet, perhaps even in a different *scriptorium*, before being inserted in a more carefully executed way than the insertion of grouping ***D***, since the text at the end of the previous grouping (*Messe*) was not simply left as a fragment once the quire labelled ~~m~~ was removed. This indicates that *Messe* was considered important to the recompilator in a way that *Trois mortes* was not, and helps us better comprehend the reason for the inclusion of *Lucidaires* at this moment in the codex. Perhaps at the instruction of a potential buyer, the recompilator decided to insert *Lucidaires*, a prose text considering various matters of faith and Church doctrine, immediately after *Messe*, another prose text discussing the workings of church life. The production and insertion of grouping ***F*** was intended to make sure that both texts could be read one to the next, highlighting the significance of both texts in the minds of the recompilator and, possibly, the first owner of the codex once it left the *scriptorium*.

While there are elements of the early life of fr24432 about which we can never be sure, our in-depth codicological examination has revealed that it did not have a straightforward birth. Instead, the codex is an intriguing kind of composite manuscript: it consists of a number of once distinct manuscript units, but the bulk of these appear certain to have originated from the same workshop, and they were all brought together into a single codex before the manuscript left its *scriptorium*. On this basis, let us now turn to consideration of a number of broader questions, concerning the manuscript's localisation, ownership, and dating.

[32] These are all also, as we shall argue in Chapter 3, the work of the same author, indicating that these texts were of particular interest to the manuscript's compiler, or its original patron.

1.3 Localisation

We have established that, despite undergoing several periods of change, fr24432 likely originated within a single *scriptorium*. There is unfortunately no clear evidence that would allow us to identify exactly which *scriptorium* this would be. Nevertheless, there are certain clues which can be used to indicate the likely general area of compilation. Rouse & Rouse (2000, ii: 60) claim that the manuscript was produced in Paris, while Busby (2002, ii: 578) suggests that it was produced a little further north, in the Oise, and Azzam *et al.* (2010: 16) suggest that both options are possible. Both the contents and language of the texts in fr24432 can shed further light on the matter.

There are two texts which make specific and repeated reference to locations in the Île-de-France, both of which appear within the base manuscript unit, such that we may assume the manuscript's original compilators and earliest readers at least had an interest in the areas discussed, and may have had more intimate knowledge. In Guillot de Paris' *Le dit des rues de Paris* (ff. 257va–261^{vb}), the author-narrator gives us a guided tour as he eats, drinks, and whores his way around the city. The focus here is hyper-local: individual streets and even specific buildings are mentioned in a way that invites readers to compare the narrator's experience with their own, such that the text could have been of particular interest to those who knew Paris' urban geography well.

> E*ncon*tre est la Rue Judas
> P*ar* la rue du petit four
> *C'on* apele le Petit-Four ;
> S*aint* Ylaire & p*uis* Clos Burniau,
> Ou l'en a rosti mai*n*t Bruliau.
> Puis la Rue du Noier
> Ou pluseurs dames p*our* louier
> Font souvent batre leur cartiers. (f. 258^{rb})

Further evidence for a focus in and around Paris can be found in the very next item in the codex, *Lendit* (ff. 261^{vb}–262^{vb}). Here the narrator offers verbal advertisement through a tour of

the *foire du Lendit*, a market that took place every year in June on the Saint-Denis plain, halfway between Paris and Saint-Denis (Franklin 1874: 12). Its position after *Rues* highlights the connection to Paris, and suggests that the patron of the original base manuscript may have held an interest in Paris which was predominantly mercantile.

That said, *Lendit* is notable also in that it mentions a large number of locations from which the merchants at the fair are said to have journeyed, of which some are far-off lands but many are also centred around northern France:

> Je les doi b*ie*n no*m*mer *com* celles
> Q*ui* plus belles son[t] a voir,
> Ce vo*us* fai-je b*ie*n assavoir,
> Cambrai cité & Moncornet,
> Mauberge & Aves i met,
> Nogent le Rotro & Dinein,
> Manneval, Torot & Caen,
> ... (f. 262va)

It is possible, then, that the inclusion of texts with a focus on northern Paris and its surroundings comes as a result of the importance of such locations to the whole of northern France, suggesting that fr24432 could have originated somewhere in these regions. This is apparently supported by the third reference we will consider here, which is to the town of Saint-Quentin, now part of the Aisne department, which is mentioned twice in the manuscript, but has much greater significance than this relative paucity of reference would appear to suggest.

The first reference to the town within the codex is made on f. 115ra within C*hevalier et escuier*, in which the author names himself as 'Jehan de Saint-Quentin.' On its own, this is not a particularly remarkable reference, and as Kedar (1973) has noted, medieval toponymic bynames are notoriously unreliable as a method of determining a person's place of origin. Jehan de Saint-Quentin is nevertheless significant to fr24432 in that modern scholarship has

assigned a further 21 texts within the manuscript to the same author,[33] and in Chapter 3 we will assign an additional 2, again both within fr24432.[34]

Within this particular manuscript, then, this otherwise unknown individual plays a surprisingly significant role. Given the lack of explicit attributions included in these other texts, it is certainly possible that such associations between the texts and Jehan de Saint-Quentin were unknown or uninteresting to the scribes and earliest readers, but what is notable for our purposes here is that the poems attributed to him are largely homogenous in form, style, and inspiration, and are almost all unique to fr24432.[35] It is therefore possible that the popularity of these texts was local, with the work of Jehan de Saint-Quentin or those like him known and enjoyed in a single localised area, perhaps as far north as Saint-Quentin, but perhaps more likely, given the presence of other texts with an explicitly Parisian focus, a little closer to Paris.

The second reference to Saint-Quentin within the codex reinforces our suggestion that the manuscript's compilation, and indeed the composition of many of its texts, did not take place as far north as this town, but that there are nevertheless geographic connections outside of the place of production. This reference appears within a series of short songs described in the *incipit* as *Serventois a Valenchiennes* (ff. 303ra). The very incipit indicates an association with northern France, referencing a town which is near the modern-day border between France and Belgium known for its literary competitions (Hécart 1833: 1–2), and many of the poems are described as 'couronné,' that is, winners of such a competition. The final line of the last poem references Saint-Quentin (f. 310^{vb}) as a place frequented by the narrator, indicating that this poem may have been composed in or around that town, and this is supported by the

[33] Most recently, see Munk Olsen (1978). The texts in question appear at items 3, 11, 20, 28, 30–42, 53, 55–57, 61, and 74.
[34] *Guillaume* and *Robert*, which appear at items 1 and 54.
[35] Again, for further details on this author and his work, see Chapter 3.

linguistic evidence, since the *scripta* of this poem, and indeed those across the whole of the *Serventois*, points towards far northern origins which are distinct from the rest of the codex.

While there are a variety of dialectal features present across the language of fr24432, most of the manuscript, including those texts attributed to Jehan de Saint-Quentin, is written in a relatively conventional Franco-Picard *scripta*. The poems of the *Serventois*, by contrast, stand out as being considerably more Picard.[36] The possessive adjective *sen* for central *son* (e.g. f. 304^{rb}) is a clearly Picard feature (Gossen 1976: 31), and spellings such as *ynfier* (f. 303^{rb}) for central *enfer* reveal the Picardian feature of the diphthongisation of tonic checked /ɛ/ (Gossen 1976: 59), as well as the presence of ⟨i⟩ and ⟨y⟩ before a nasal consonant where in more central *scriptae* one might expect ⟨e⟩. The spelling *boine* (e.g. f. 304^{vb}) displays the Picard tendency to diphthongise tonic /o/ before a nasal consonant (Gossen 1976: 83), while the demonstrative pronoun *chiaus* (e.g. f. 304^{rb}) is particularly clear, displaying the Picard triphthongisation of /e/ (Gossen 1976: 62–63), alongside the use of ⟨ch⟩ for central ⟨c⟩ /s/ which is word-initial or syllable-initial and following a consonant (Gossen 1976: 91. See also *merchi*, *che*, *enfanche*, and *fianche*, all within a single stanza on f. 304^{vb}).

The use of ⟨k⟩ in place of the central ⟨qu⟩ is also particularly noticeable, with the *Serventois* containing 203 of the 245 instances of the character across the whole of fr24432, most of which are for the words *ke* (126 instances) and *ki* (54). There are nevertheless still 164 uses of ⟨qu⟩, which again are largely split between *que* (65) and *qui* (31), but as Gossen (1976: 95) notes, this does not make the scripta any less Picard, since both methods are used by Picard scribes interchangeably. There is also the possibility here of some scribal intervention, be that before or during the compilation of fr24432, which may have occurred as the text was transmitted away from its original place of composition, into areas with more conventional

[36] Here I refer to the features of Picard *scriptae* identified by Gossen (1976).

scriptae, and copied by scribes for whom such spellings were unfamiliar. The fact that any intervention is not total, however, reveals that these scribes and their readers were apparently comfortable with even the most obvious elements of the Picard *scripta*.

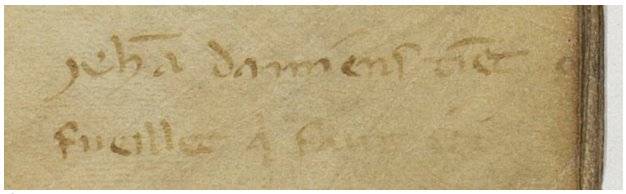
The triphthong ⟨ieu⟩, meanwhile, has not been reduced to ⟨eu⟩ in words such as *dieu* (e.g. f. 303ra) as we might expect from a Picard *scripta*, although this is less common for texts written around Saint-Quentin (Gossen 1976: 55), and it is once again possible that scribal regularisation played a role here.

On the whole, it is therefore clear that the poems of the *Serventois* are distinct amongst the texts of fr24432 in their use of clear Picard features. It is this exception that proves the rule, such that the manuscript itself is unlikely to have been compiled as far north as Saint-Quentin. The short poems that had seen success within the Valenciennes *puy* appear to have been more widely read outside of their immediate *loci* of composition, such that their presence in fr24432 does not suggest it was produced in the immediately surrounding area.

It does not necessarily follow either, however, that the manuscript must have been produced within the Île-de-France. Instead, the decision to include a text like *Serventois* with strongly Picard features suggests that the manuscript came into existence somewhere between the Île-de-France and Picardy, and was influenced by traditions and texts in both areas. Busby's (2002, ii: 578) suggestion that the manuscript may have been compiled in Oise is therefore more appealing, since it offers a location between the extremes of Saint-Quentin and the Île-de-France, appearing a reasonable hypothesis in the absence of more specific evidence.

One final detail is worthy of consideration here. Rouse & Rouse (2000, ii: 60) observe a marginal note on f. 96^{r}, referencing 'Jehan d'Amiens.' This hand is roughly contemporary with the production of fr24432, and appears to reference a missing page of the exemplar from which this text (an alexandrine version of *La vision de saint Paul*) was likely copied.

Figure 1.25 – a marginal note concerning one Jehan d'Amiens.



Jeha*n* d'Amiens tie*n*t c[?]
fueillet q*ui* faut ici

f. 96rb

Given that the text on this and the surrounding folios runs uninterrupted with no apparent material or scribal break, this note was therefore most likely added while the manuscript was still in the early stages of production in its *scriptorium*, with Jehan d'Amiens an individual involved in manuscript production, although most likely not with the production of fr24432.

Again, we ought to be cautious about reading too much into the byname of this particular Jehan, but the association with northern France is notable. On the other hand, and again reinforcing the sense that fr24432 straddles both Parisian and more northern influences, Rouse & Rouse (2000, ii: 60) suggest that this Jehan d'Amiens may have been a kinsman of another Jehan d'Amiens, known to have been in Paris in 1298, likely a half century after the production of fr24432. While a totally unverifiable claim, it is nevertheless a tantalising one for the localisation of our manuscript, since the later Jehan d'Amiens is recorded as having worked on the Rue aux Fèves, a street on the Île de la Cité which also appears in the *Rues*:

> *Et* me samble que l'autre chief
> Descent droit en la rue a Feves. (f. 259ra)

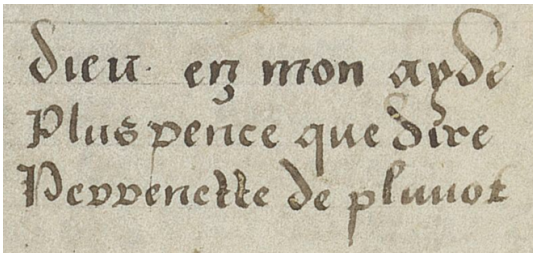
Fr24432 is thus a manuscript in which a number of geographical influences can be seen, ranging from central Paris to what are now the most northern reaches of mainland France. It appears most likely that the manuscript was compiled somewhere between these two extremes, with Busby's suggestion of Oise appearing particularly plausible, even if an exact localisation is impossible.

1.4 Ownership

Very little extra- or intra-codicological evidence survives of the manuscript's early ownership. There is no indication of the original patron or intended readership for the manuscript, and very little can be known about the manuscript's ownership before the sixteenth century, when two hands make contributions to its contents and sign their work, pointing to their involvement in the manuscript's cycle of ownership across the centuries.[37]

Perrenette de Pluvot, notably an apparently female owner, contributes two additions to the codex. Nothing is known of her biography, but we note that the first contribution appears to consist of two mottos, perhaps associated with her family, whereas the second is a copy of Clément Marot's *Chanson XXII*, originally published as part of his *Adolescence Clémentine*, which proved particularly popular in aristocratic Paris circles.[38] This could suggest that around this time de Pluvot was associated with the Paris court, and had possession of the manuscript in the 1530s, around the peak of Marot's popularity.

Figure 1.26 – Perrenette de Pluvot's contributions to the codex.

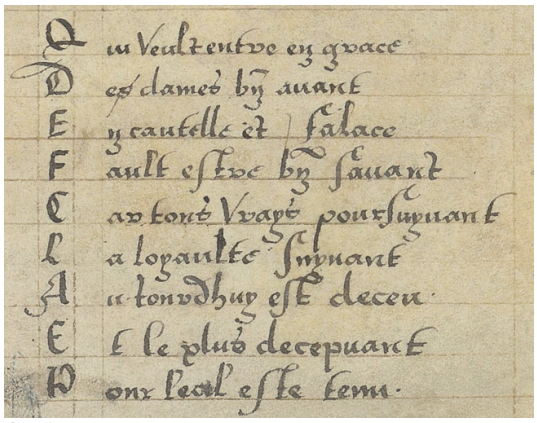


Dieu en mon ayde :
Plus pence que dire.
Perrenette de Pluvot

[37] Grange (2018) notes the presence of the name 'Adam Raymont' on f. 279^{rb}, who is suggested as a possible owner. In fact, this is a mistaken attempt by a later reader to identify the author behind item 66 (Raimont Badaut), who hides his name in an acrostic within the final lines of the poem. This reader clearly has an interest in authorial attribution, appearing again on ff. 388ra and 390^{rb} to assert that both items 79 and 80 are 'Par Watriquez.'

[38] Saintsbury (1911). Many early editions of the text exist, and Saintsbury puts the first publication date at 1532. Here, I cite a 1533 edition, in which the poem in question can be found on p. 85.

f. 265rb



f. 407rb

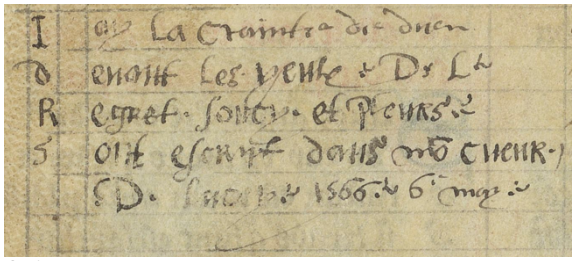
Qui veult entre[r] en grace
Des dames, b*ie*n avant
En cautelle et falace
Fault estre b*ie*n savant.
Car tout vraye poursuyvant,
La loyaulté suyvant,
Au jourd'huy est deceu.
Et le plus decepvant
Pour leal est(e) tenu.

The first insertion by de Pluvot is likely a mark of ownership, but the second is seemingly evidence that the manuscript was still considered an evolving document nearly 200 years after it was first compiled. De Pluvot evidently considered it appropriate to add a snippet of poetry that resonated with her in amongst much older poems that had a similar impact, suggesting that her interaction with the codex was focussed as much on its intertexts as its texts, and offering particularly striking evidence of the intra-codicological thematic networks that the manuscript likely elicited in its earlier readers.

A second sixteenth-century owner by the name of Denis Lucey contributes two very similar additions which appear near to de Pluvot's, of which one (f. 279va) is inserted between two texts, while the second (f. 402va) appears within a gap apparently left for illustration within *Prise*, a verse allegory in which the narrator is the object of a hunt by Love and his entourage. Since Lucey helpfully signs both of his contributions as having being completed on the 6th of May 1566, we can assume the manuscript was in his possession by that point, most likely after it had been owned by de Pluvot, given the early-sixteenth-century style of her hand. Both Lucey's contributions appear within a few folios of those by de Pluvot, suggesting that Lucey, if indeed he was the later of the two owners, made his mark on the codex in a way that was

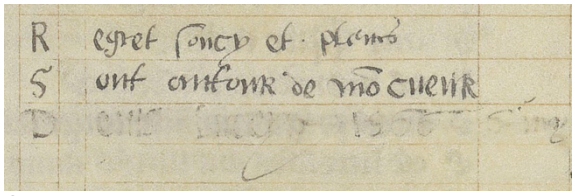
intended to supplant de Pluvot as the previous owner. As a result, we may assume that these notes were added shortly after he acquired the codex.

Figure 1.27 – Denis Lucey's contributions to the codex.



f. 279va

J'ay la crainte de Dieu
Devant les yeulx.[39]
Regret, soucy, et pleurs
Sont escript dans mo*n* cueur.
D. Lucey, 1566, 6e may



f. 402va

Regret, soucy, et pleurs
Sont autour de mo*n* cueur
[Denis Lucey, 1566 6e may]

The position of both sets of additions within the manuscript could also indicate which texts were of greatest interest to a sixteenth-century reader. Lucey and de Pluvot's first contributions appear either side of *Arbre*, while their second contributions appear in and around *Prise* and *Mesdisans 2*, all three of which are long allegories on matters of love, with themes that would fit with the melancholy nature of the contributions. Given that these positions in the manuscript are not prominent simply by dint of being near to the beginning or the end of the codex, it would therefore appear that Lucey and de Pluvot felt that these items were of particular significance and turned to them regularly, or expected others to do so. As a result, we may hypothesise that much of the sixteenth-century interest in fr24432 centred primarily around such love allegories.

Notably, a later owner appears to have scratched out Lucey's signature on f. 402va. We may hypothesise that this owner could well have been the next possessor of the codex after Lucey,

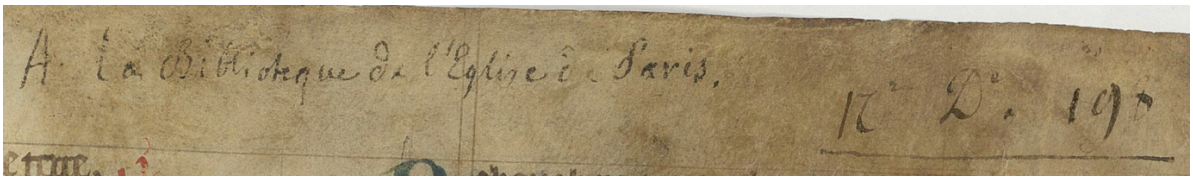
[39] The ':- D :- L :-' on this line appears to simply be an attempt by Lucey to fill the line using his initials.

since his name is erased but not de Pluvot's. This new owner also most likely gained ownership through purchase or exchange rather than inheritance or gifting, since it seems unlikely that they would have wanted to erase the signature of a relative or a friend.

In BnF Arsenal 4629, an eighteenth-century manuscript composed by Étienne Barbazan (1696–1770), entitled 'Catalogue des manuscrits de l'Église de Paris', we find the first external reference to fr24432. The manuscript is described as a 'Recueil de 80 pieces en vers françois qui sont Romans, Satires, dits Moraux, Contes, quelques pieces en prose' (214). Barbazan informs us that the manuscript was once in the collection of a M. Chevane, likely Jacques-Auguste de Chevanes (1624–1690), an 'Avocat au Parlement de Dijon [...] fort conu dans la litterature' (141), who had amassed a collection of some seventy manuscripts, and who is acknowledged in other near-contemporary sources as having a notable 'connoissance des anciens titres' (Bruys 1751: 16).

Chevane's collection passes to a nephew after his death,[40] who in 1752 sends it to a member of the Collège de Cambrai within the University of Paris to arrange for its sale. It is seemingly Barbazan who completes the sale process, selling part of the collection, including fr24432, to Abbot Jean Omer Joly de Fleury (1700–1756) on behalf of the library of Notre-Dame de Paris. There, it receives an *ex libris* and the shelf mark N.-D. 198, which are both added to f. 0r.

Figure 1.28 – ex libris *and shelf mark added to f. 0r after the 1752 sale of the codex.*



[40] Grange (2018) suggests a 'M. Dilan' as the recipient of this bequest, a name found in other manuscripts once in the possession of de Chevanes such as BnF NAL 1746. Older accounts such as that of Papillon (1742: 142) and Bruys (1751: 17), however, suggest instead a 'M. Thomas.'

This shelf mark is retained when the manuscript is sold in April 1756 to the Bibliothèque royale, along with some 300 others, for 50,000 *livres*. This short turnaround period for the codex within the library of Notre-Dame de Paris is explained by the death of Joly de Fleury earlier in 1756, which appears to have resulted in a change in priorities at the cathedral (Franklin 1867: 40). Once in the Bibliothèque royale, the codex then gains library stamps such as that found on f. 0^{r}. Later, when the institution becomes the Bibliothèque impériale, further stamps are added, such as that on f. 57ra.

Figure 1.29 – examples of 18th- and 19th-century library stamps from the Bibliothèques royale (L) and impériale (R).

f. 0r *f. 57ra*

The earlier, royal librarian appears to have been more reserved in their placement of stamps, doing so outside of the text block and only at the beginning and end of the codex, whereas the later, imperial librarian is more exuberant, adding a total of eleven stamps, all of which appear inside the text block at gaps in the text, and in one instance even on a section of water damaged text on f. 35^{rb} (Figure 1.66).[41]

We therefore have little knowledge of the initial ownership of fr24432, or of the century or so between its acquisition by Denis Lucey and its appearance in the collection of Jacques-Auguste de Chevanes. What is more certain, however, is that the manuscript was in private circulation by the sixteenth century, being bought and sold by individuals with no

[41] Royal stamps appear on ff. 0^{r} and 434^{v} (which incidentally supports the hypothesis discussed above that the disappearance of material originally at the end of the codex is not a modern loss). Imperial stamps appear on ff. 35^{rb}, 57ra, 84^{rb}, 99^{rb}, 281^{vb}, 310^{vb}, 317^{rb}, 327^{v}, 328(bis)^{ra}, 385^{rb}, and 388ra.

obvious connection to each other, with the manuscript entering institutional ownership for the first time in the eighteenth century.

Working back from the established chain of acquisition, it does appear probable that fr24432 was in private hands from its earliest beginnings in the fourteenth century. Grange's (2018) suggestion that the possible original patrons were 'a rich merchant or wealthy rural family' therefore seems plausible, especially given the presence of texts we have already examined which discuss merchant activity in and around Paris. Indeed, such a hypothesis is further strengthened by the presence of ecclesiastical texts in vernacular translation,[42] suggesting an interest in theology and the workings of the Church, but no direct involvement in them.

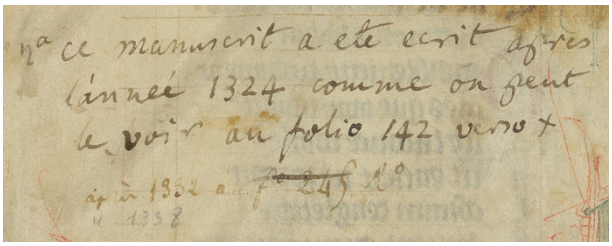
That said, given the numerous processes of recompilation and modification that we have observed for fr24432, it would be simplistic to claim that the original patrons must also have been the original owners. It is possible that the base manuscript was commissioned by a patron who was ultimately never able to take possession of it, and the later processes of recompilation could therefore have occurred to align the codex with the wishes of a later patron. If fr24432 was initially in the possession of a merchant family, this could well have been as much down to accident as to original design.

1.5 Dating

As with questions of localisation and ownership, it is possible to gain a vague idea of when fr24432 was compiled, but an exact date proves elusive. Numerous attempts have been made by the manuscript's readers across the centuries to identify the date after which it must have been compiled, as is evidenced by the marginal debate on f. 1ra (Figure 1.30).

[42] e.g. *Lucidaires*, *Salomon*, and *Misere*.

Figure 1.30 – three (probably 19th-century) readers debate the manuscript's earliest possible date of compilation.



N*ota* ce manusctiy a ete ecrit apres
l'année 1324 comme on peut
le voir au folio 142 verso x
après 1332 a[u] ~~*folio* 245~~ R*ecto*
" 1338

The origins for these marginal claims are found in the presence of a number of dates incorporated into texts as dates of composition, or otherwise by the texts' scribes in the surrounding paratext.[43] Ultimately none of these readers find the manuscript's true *terminus post quem*. The latest date in the codex is found on f. 266rb, in which *Arbre* references St. George's Day 1345, the 23rd of April. It is clear, then, that the manuscript could not have been compiled before this date, and it is unlikely that our text is particularly close in the transmission cycle to the autograph copy of *Arbre*,[44] which therefore likely pushes the date of compilation towards the second half of the fourteenth century. Given the presence of texts dating earlier in the century, however, with the earliest specific date given as 1318, as well as much earlier texts such as Marie de France's *Yonec*, likely a product of the late twelfth century (Busby, 2011), it seems that the codex was not compiled any later than the 1350s.

This hypothesis is supported by more physical and structural evidence present in the manuscript.[45] We have already noted that the ruling of the text block was carried out using a number of different materials. This variety served to demonstrate the internal material diversity of fr24432, but it can also contribute to our efforts to date the codex. Nixon (1993: 7)

[43] In the former category: 1332 (f. 245ra), 1338 (f. 17rb), and 23rd April 1345 (f. 266rb). In the latter: 1313 (f. 146va), 1318 (f. 148rb), 1320 (f. 152ra), and 1324 (f. 142vb). Items 44–47 therefore all contain a composition date in their *explicit*, suggesting both that they came from the same exemplar, and that the scribe in question (hand 6) saw no reason to standardise the paratextual format alongside other texts with no available date.
[44] See the examination of the manuscript's decoration in the appendix on p. 292.
[45] For the codicological evidence examined here, I use the elements for the dating of French manuscripts between the 12th and 14th centuries proposed by Nixon (1993: 5–11). Given an earliest possible date of composition is known, I have only examined those elements which display a difference in practice between the early and late 14th century.

notes that lead, the predominant material used for the ruling of leaves in the thirteenth and early fourteenth centuries, diminished in popularity over the middle two quarters of the fourteenth century before its dominance was replaced by that of ink ruling towards the end of the fourteenth century. That fr24432 contains examples of lead, wax crayon, and ink ruling would therefore place the manuscript's production in the middle of this period of change, that is towards the middle of the fourteenth century.

There is further evidence that pulls the dating of fr24432 towards one end of the fourteenth century or the other, suspending the codex between two extremes and making a mid-century date more likely. On the one hand, the use of pen flourishing for larger initials which is 'concentrated around the initial with short lines and bead rows framing the initial and short wavy lines and loops extending out from the initial' (Nixon 1993: 10), as in Figure 1.20, is a particular feature of the first half of the fourteenth century. On the other hand, the manuscript's quiring points towards later in the fourteenth century rather than earlier. As we have seen, the codex predominantly consists of quaternions, that is quires of 4 bifolia, or 8 folios. Busonero (2022: 230–231) has suggested that quires of 6 bifolia (senions) were predominant in the first half of the fourteenth century, and that it is only in the second half of the century that the quaternion returns to the dominant position it held in the twelfth century.

A synthesis of all the available evidence therefore points to the likelihood that the compilation of fr24432 began in the late 1340s or early 1350s, but given the processes of modification and recompilation that we have identified, it is likely that period from planning to completion took some time, perhaps several years. It is perhaps no coincidence that the dates we have identified line up with the period in which the Black Death was wreaking havoc across medieval Europe. Could it be that this explains the apparent time gaps between the copying of the base manuscript and the acts of recompilation? Was *Lucidaires* inserted in grouping G to

satisfy the demands of a new patron following the death of the original individual? Is the presence of repeated material the result of a recompilatory process combining the work of scribes who never lived to complete their work? We may never know, but knowledge of the likely dates of the manuscript's composition does paint a picture, and not a particularly pleasant one, of the lives and possible fates of those involved in the compilation process.

Our examination of the codex has revealed a number of hitherto unknown details, but it is far from comprehensive, and there is much about the production of fr24432 that remains inaccessible, at least for the moment. It is therefore easy to agree with Étienne Barbazan, who in his catalogue described the codex as 'tres curieux' (214). In the following chapters, we will examine these curiosities further, taking two case studies for which the groundwork has already been laid in this chapter. In the first, we will examine the role that an analysis of abbreviation can play in the identification of scribes and other features of a manuscript. We will then turn to questions of the attribution of authorship, examining the texts found both within fr24432 and outside it which may be considered the work of a single author.

2. Sc[i]bal abb'viatiō

Scribal abbreviation and its uses in fr24432

2.1 The study of medieval French manuscript abbreviation

Abbreviation is near ubiquitous in Western medieval manuscripts, and yet until relatively recently it has received little scholarly attention. For decades, it has been widely dismissed as a feature which poses an intellectual challenge for the scholarly editor, but which is otherwise of minimal analytic interest, having existed solely as a way for a scribe to save time, space, or writing material.[1] In many instances, the presence of abbreviation can indeed be explained by these factors, when there is evidence that scribes have been constrained by the limitations of a particular *mise en page*, having to abbreviate words in order for a line of text to fit the format, or when the demands of a particular copying project require sparing use of ink or parchment.

That said, recent scholarship has begun to acknowledge that practical factors are far from the only explanation for the presence of abbreviation in manuscripts.[2] Geneviève Hasenohr (2002: 91), for example, notes the presence of abbreviations in lines of text that would never have filled the available space even if left unabbreviated, as well as the fact that certain modes of abbreviation, such as the use of superscript letters, are not a particularly effective way of saving time or ink. Indeed, even in the most lavishly decorated codices, where aesthetic considerations win out time and again over the preservation of resources, we still find abbreviations occurring with remarkable frequency.

There is therefore much more to manuscript abbreviation than the traditional account would suggest: it cannot be explained solely by practical factors, and instead we must take a broader approach. This chapter attempts to do exactly that, building on recent work to consider scribal abbreviation within medieval French manuscripts in particular, through an examination of fr24432.[3] In doing so, I attempt to display the utility of the statistical analysis

[1] See, for example, Cappelli (1982: i).
[2] For an excellent overview of the state of the field regarding medieval abbreviation, see Honkapohja (2021).
[3] I use the term 'medieval French' here in a broad sense. Linguistically speaking, many of the texts found within fr24432 could fall under the definition of 'Middle French,' while others could be categorised as 'Old French.'

of abbreviation as an addition to the manuscript scholar's toolbox, which can be used to reinforce conclusions drawn from more traditional methods of study, and open up new avenues for research.

We begin the chapter with an examination of certain theoretical considerations pertaining to the study of manuscript abbreviation, particularly as relevant to medieval French. Following this, we put this theory into practice through a case study on the use of abbreviations in fr24432.

2.2 Theoretical considerations

For the most part, existing scholarship has classified manuscript abbreviation according to formal typologies. That is, the existing typologies have been concerned first and foremost with categorising the signs and symbols used in the process of abbreviation, rather than the words or characters which they represent. While this formal approach is common, and certainly has descriptive utility, I will argue here that a more functional approach offers greater possibilities for analytic exploitation.

2.2.1 Formal typologies

In his 1890 *Manuel de paléographie latine et française*, perhaps the most comprehensive amongst the nineteenth-century palaeographic *magna opera* in its examination of medieval French, Maurice Prou separates methods of abbreviation into five principal categories, which still appear in more recent scholarship and which we examine briefly here.

Abbreviation by the use of initials (1), as in ⟨RIP⟩ for ⟨requiescat in pace⟩, is a simple enough concept, but is notably a mode of abbreviation common in Latin and yet almost entirely absent in medieval French. The occasional exceptions involve the abbreviation of personal names which appear repeatedly within a given text, such as the abbreviation of ⟨Gauvain⟩ to

⟨G.⟩ found in Arthurian romance (Lepage 2001: 108). Abbreviation by contraction (2), meanwhile, through which two or more characters from within a word are retained while others are elided, is a more common feature in medieval French manuscripts. These forms are often borrowed directly from Latin, as in ⟨bn̄⟩, which can be expanded to the two cognates ⟨bene⟩ in Latin and ⟨bien⟩ in medieval French, but there are also some examples of abbreviation by contraction appearing independently in medieval French, for example ⟨ch'r⟩, a common abbreviated form for ⟨chevalier⟩ 'knight'. Abbreviation involving the use of superscript letters (3) is more common still, both for function words like in ⟨q^{i}⟩ for ⟨qui⟩ (present in Latin and medieval French), and in more contentful words like the name ⟨Alixandre⟩, which can appear abbreviated to ⟨Alixand^{e}⟩. Abbreviation by suspension (4), meanwhile, involves shortening a word to just its first few letters, a type of abbreviation which is widespread in Latin, where forms such as ⟨īt⟩ for ⟨item⟩ are common, but which appears far less frequently in medieval French, almost to the point of absence, again with the exception of a few personal names like ⟨Lan.⟩ for ⟨Lancelot⟩. One could argue, however, that the use of superscript characters to indicate abbreviation at the ends of words, as in ⟨q^{i}⟩ above, constitutes a kind of abbreviation by suspension. Finally, the category of special and conventional signs in Prou's typology (5) provides the greatest number of examples in medieval French manuscripts. These signs include the macron bar ⟨–⟩ atop a character to represent a following nasal consonant, the sign ⟨9⟩ which can be used to represent ⟨us⟩, amongst other sequences, and the symbol ⟨7⟩, the so-called Tironian note, which represents ⟨et⟩ ('and' in both Latin and medieval French).[4]

While Prou's terminology, drawn primarily from Latin palaeography but applied to the medieval French context, is certainly useful for descriptive purposes, it does not appear that

[4] In some scriptae, such as Anglo-Norman, the expanded form for the Tironian note can be ⟨e⟩.

his typology of abbreviation contains categories which are sufficiently discrete to be useful for analytical purposes. Taking the example of ⟨v^{9}⟩, a common abbreviation for ⟨vos⟩, ⟨vous⟩, or ⟨vus⟩, we find a form which could be classed as an example of abbreviation by superscript letters, suspension, and special signs, thus cutting across three of Prou's types. Further, we may also choose to question why, for a study of abbreviatory practices within manuscripts, it would matter whether one character from the beginning of a word is retained (as in abbreviation by initials) or more than one (by suspension). Similarly, there is limited analytic value in distinguishing between superscript and non-superscript signs, since in many cases such distinctions appear to exist on a spectrum, with special signs such as ⟨9⟩ appearing in both superscript and in-line forms, and frequently somewhere between the two extremes, making discrete classification difficult. Indeed, even those characters which are unambiguously superscript often take a form distinct from their non-superscript versions, as often occurs in Gothic scripts with ⟨^{a}⟩, meaning that it can be unclear to what extent a scribe producing both forms would have treated them as versions of the same character.

Nicolas Mazziotta (2008) offers a counterpoint to Prou's perspective, producing a typology which is distinct from the traditional account, and yet still takes a primarily formal approach, focussing on the details and composition of the symbols used to mark abbreviation in medieval French manuscripts.

Grammèmes, the root node in Mazziotta's typology, are all discrete written forms, whether or not they are associated with abbreviatory practices. These are divided into two sub-types, the first of which are the *plérégrammes*, forms which carry a particular associated meaning, be that more phonological (for *caténogrammes*) or more lexical (for *logogrammes*). Outside of the abbreviatory system, most forms appearing in manuscripts written in medieval French are therefore *caténogrammes*. The second sub-type of *grammèmes* are the *cénégrammes*, forms with a

distinctive, modifying function, which may appear in the same linear axis as surrounding *plérégrammes* (for *linéogrammes*), or in some other position such as above or below the line (for *périgrammes*). Abbreviation markers in medieval French texts therefore occupy three of the four third-level categories: whole-word abbreviations like ⟨7⟩ for ⟨et⟩ are *logogrammes*, the intra-linear variant of ⟨9⟩ is a *linéogramme*, and the macron bar used in the abbreviation of nasal consonants is a *périgramme*.

There are certainly criticisms to be made of Mazziotta's approach. Firstly, as Dominique Stutzmann has noted (2010: 266), it does not offer a satisfactory account of abbreviation by contraction, and it also results in the apparently arbitrary separation of what may be considered one symbol into more than one. For example, the bar which crosses the descender in ⟨p̷⟩, used in medieval French, as in Latin, to abbreviate ⟨par⟩, ⟨per⟩, or ⟨por⟩, is treated by Mazziotta's typology as a separate symbol, a *périgramme* to be precise. However, this does not match the reality of its usage in medieval French, since this horizontal line on the descender of a character is usually only found on a ⟨p⟩, and thus does not appear to be an independent combining character, rather a constituent part of an alternate version of an existing character.

Even in the case of unambiguous modifying characters (*cénégrammes*), if our goal is the analysis of abbreviation use in medieval French, rather than its description, then Mazziotta's typology can also be subject to the same criticism that we raised for Prou's account regarding the distinction between superscript and non-superscript abbreviation markers. As with Prou's typology, there appears to be no analytic benefit gained through Mazziotta's separation between modifying signs that appear on the same spatial axis as other characters (the *linéogrammes*) and those which do not (the *périgrammes*), given that this is likely just an accident of tradition.

While valuable as palaeographic descriptors, then, the existing formal typologies of abbreviation do little to take us any closer to being able to gain an understanding of the process of abbreviation as experienced by scribes and their earliest readers. An analysis which focusses solely on the symbols used for abbreviation cannot offer a meaningful explanation for why such abbreviation occurred in the first place. In order to do that, we must also look a little more abstractly, moving from a formal typological approach to a functional one, considering the specific mechanisms by which words or characters are abbreviated rather than the symbols used to represent them.

2.2.2 Functional typology

Grapheme- and lexeme-specific abbreviation

In a brief yet incisive presentation of approaches to manuscript transcription, Matthew Driscoll (2006: 259) makes a critical contribution to the study of manuscript abbreviation, suggesting that 'a case could be made for distinguishing between abbreviations with a lexical reference (suspensions, contractions, and a number of brevigraphs) and those with a graphemic reference (superscript letters and signs and the remainder of the brevigraphs).'[5] While Driscoll still makes reference to the formal typological distinctions whose utility I have already disputed, the shift in focus from form to function which his typology encourages is one which offers a great deal to the analysis of abbreviation in medieval French. It is through a functional lens, I will argue, that the medieval French abbreviatory system can best be understood.

Abbreviation with graphemic reference, which I refer to here as *grapheme-specific abbreviation*, involves the use of an abbreviatory mark to indicate the abbreviation of a set sequence of graphemes commonly associated with that particular mark, regardless of the word in which

[5] By 'brevigraphs,' we understand Prou's special and conventional signs.

the abbreviatory mark appears. An individual who knows the referents of a given grapheme-specific mark can therefore expand any word containing it, without needing to draw on a great deal of lexical knowledge. An example of grapheme-specific abbreviation is therefore the macron bar, appearing above a character to indicate a following nasal consonant, which can be expanded by a reader or used by a scribe, even if they have never before encountered the word in question.

Lexeme-specific abbreviation, meanwhile, involves the use of a standardised mark or sequence of marks to represent an individual word form, which may bear greater or lesser resemblance to the word's unabbreviated form, and which cannot simply be expanded through knowledge of grapheme-specific rules alone. That is, lexeme-specific abbreviation is where a symbol or sequence of symbols refers to a full lexeme, with the connection between the two either learnt or inferred from lexical knowledge.

Some abbreviatory symbols of this type may be fully logographic, in that they refer to a full word through a single symbol which bears no relation to the lexeme's unabbreviated form. There are relatively few distinct forms of this type in medieval French, but they include the Tironian note ⟨7⟩ for ⟨et⟩ 'and', as well as the symbol ⟨÷⟩, representing the third person singular present tense indicative form of the verb 'to be', both of which are observed in Latin abbreviation and are therefore to some extent language independent, reinforcing their lexical, rather than graphemic, reference.[6] On the basis of Roger Lass and Margaret Laing's description of such symbols as functioning as '"pure" logographs' (2013: §2.2.1), we may call these 'pure' logographic abbreviations.

Other examples of lexeme-specific abbreviation involve the use of abbreviatory symbols alongside non-abbreviatory ones in ways that resemble grapheme-specific abbreviation, with

[6] The form 'est' can sometimes also be abbreviated as ⟨ē⟩, but there are very limited examples of this in fr24432.

the difference being that the abbreviated word cannot be understood solely through knowledge of the standard grapheme sequences referenced by individual abbreviatory symbols. These, again based on Lass and Laing's categorisation, are what we may refer to as 'impure' logographic abbreviations, since they contain graphemic clues as to the lexeme they represent, but these alone are still not enough to enable confident comprehension without knowledge of the processes of lexeme-specific abbreviation.

Examples of this sub-type are also unusual in medieval French texts, but this time they are not limited solely to abbreviations already used in Latin. One of the most frequent, for example, is ⟨ch'r⟩, a form which, as we have seen, is used to abbreviate the word ⟨chevalier⟩ 'knight,' through suspension of various internal characters. While this abbreviation is less logographic than the 'pure' examples, it is still an instance of lexeme-specific abbreviation since it can only be understood by a reader with prior experience of the form, or sufficient lexical and contextual knowledge to make an educated guess.

We should note here that grapheme-specific abbreviation in medieval French is always linear, whereas lexeme-specific abbreviation can be non-linear. For example, in ⟨Jhrl'm⟩, a lexeme-specific abbreviation of the 'impure' type, used for ⟨Jherusalem⟩, the single abbreviatory marker present in the word is used to represent no fewer than four individual grapheme sequences.[7] A grapheme-specific approach, by contrast, would require the use of multiple markers in order to achieve the abbreviation of more than one sequence, since such markers can only ever refer to a single linear grapheme sequence.

[7] This is an example of the so-called *nomina sacra*, abbreviated forms of divine names and titles which almost exclusively fall under the category of 'impure' lexeme-specific abbreviations, and which are particularly common in Latin.

Objections to the functional typology

One objection to this new typology may come from the only existing medieval treatise focussed solely on abbreviation, the fifteenth-century *Quedam regule scribendi et sillabicandi bonos libros*, which offers a number of rules laying out what constitutes 'good' abbreviatory practice, according to the individual who wrote it. The rules within are concerned only with Latin, but refer to abbreviations which also occur in the vernacular languages, including medieval French. The treatise sets out a number of contexts in which abbreviation is said to be counter to proper writing style, in particular when it would stray beyond the boundaries of a single syllable.[8] This focus on the syllable, a phonological unit, rather than to individual letter forms, may therefore reasonably lead us to question whether it would be better to think of grapheme-specific abbreviations as phoneme-specific abbreviations, with abbreviatory marks standing for sub-syllable features such as phonemes.

Fundamentally, there are two distinct types of sequence to which partial abbreviation within a word may refer: either an abbreviatory symbol refers to a sequence of sounds, or it refers to a sequence of symbols which in turn refer either to a word or to a sequence of sounds. The evidence, it appears, is in favour of the latter approach, that is taking graphemes as the units upon which abbreviation acts, rather than phonemes. This is because the graphemes to which an abbreviatory symbol may be said to refer are far more consistent than the equivalent phonological referents. A clear example of this can be found for the macron bar atop a vowel, which as we have seen is generally expanded by modern editors as either ⟨m⟩ or ⟨n⟩, but for which the set of possible phonological referents would be a good deal more complex.[9]

[8] For more on this treatise and a summary of its contents, see Bozzolo et al. (1990). It is also discussed by Hasenohr (2002: 89–90), who questions the validity of such a syllabic approach, given that in practice abbreviations so often span multiple syllables.

[9] The macron bar can also be used in lexeme-specific abbreviation to represent something other than a nasal consonant, for example atop a ⟨q⟩ to represent ⟨que⟩, or atop ⟨e⟩ to represent either ⟨est⟩. Arguably, this multiplicity of uses across different kinds of abbreviation reinforces the value of such a division.

To take a specific case, the word ⟨comment⟩ 'how' is frequently abbreviated in medieval French texts as ⟨cōmēt⟩, with this form appearing 20 times in fr24432. Such an example raises issues for the phonological approach when we think about how this word would actually have sounded. By the time fr24432 was compiled, the process of the nasalisation of vowels and the accompanying elision of nasal consonants was complete in French, and so in most dialects the word would have sounded much as it does today, with a phonological sequence somewhat akin to /kɔmɑ̃/ (Marchello-Nizia *et al.* 2020: §367, 386). If the macron bars were to be taken as having phonological reference, then in the first instance they would refer to a full phoneme, and in the second to the phonological feature of nasality. This is not impossible, but it introduces a level of complexity to the abbreviatory system that appears entirely unwarranted, especially since the graphemic approach can offer a simpler explanation. From a grapheme-based perspective, the macron simply stands for a written nasal consonant, the choice of which can be determined by graphotactic factors, as well as the reader's lexical knowledge.

There is, of course, the argument that the expansion of these abbreviations, even if it is fundamentally graphemic, is in some way phonologically conditioned, since the usual expansion of the first macron bar as an ⟨m⟩ could be said to be determined by the presence of a following bilabial consonant, while in the second case the ⟨n⟩ would be the most likely choice given the presence of the following alveolar consonant. Yet, the fact that the final ⟨t⟩ does not usually have a corresponding phonological unit points to the fact that this is a case of graphotactic, rather than phonotactic conditioning, whereby given a choice between the characters ⟨m⟩ and ⟨n⟩, ⟨m⟩ is the more likely choice before another ⟨m⟩, while ⟨n⟩ is most likely before ⟨t⟩. These graphotactic constraints certainly have their origins in the phonology, but phonological considerations are not at play here. Instead, in the context of medieval French, we may argue that, whether or not scribes felt constrained by the boundaries of the

syllable when abbreviating words, abbreviations which only form part of a word have solely graphemic reference.

2.2.3 Processes of use and interpretation

This typological division between lexeme- and grapheme-specific abbreviation therefore allows us to identify two abbreviatory modes, one focussed on the abbreviation of a whole lexeme, the other of a sequence of graphemes of at least length one. When interpreted by a reader, these two modes would appear to follow two separate pathways.

The process for lexeme-specific abbreviation is quite simple: the abbreviatory symbol represents a lexeme. This may be in the same way that we see with logographic writing systems like that of Chinese characters or Japanese *kanji*, for the so-called 'pure' type, or the word form may contain alphabetic forms alongside any abbreviatory symbols, in the so-called 'impure' type. 'Pure' lexeme-specific abbreviation usually occurs either on non-inflecting words (as in ⟨7⟩), or on specific pre-inflected forms (as in ⟨÷⟩), whereas 'impure' instances of lexeme-specific abbreviation can generally undergo inflection or derivation. Thus ⟨ch'r⟩ can inflect to ⟨ch'rs⟩ in the nominative singular or the oblique plural, just as the unabbreviated form does, and the same abbreviation can also be applied to derived forms of the same word, as in ⟨ch'rie⟩, which appears as an abbreviated form of ⟨chevalerie⟩ 'chivalry, cavalry'.

The process from grapheme-specific abbreviation to lexeme is a little more complex. In this case, the abbreviatory symbol represents a sequence of graphemes. This sequence fits into a wider graphemic structure, which itself is used to represent a given word form, and may bear greater or lesser resemblance to a word's phonological form. This process is therefore comparable to that followed by non-abbreviatory alphabetic symbols, which themselves are combined into a graphemic structure for each word alongside those graphemes represented by any abbreviation.

All this, of course, is from the perspective of a reader. Should we be interested in the perspective of a scribe producing abbreviated forms, then we simply work the other way, from a lexical form which is associated either with a lexeme-specific abbreviation or with a phonemic structure, with the latter represented by a graphemic structure which in turn, when broken down into individual graphemes, may be abbreviated using grapheme-specific abbreviation.

From this scribal perspective, a further distinction between grapheme- and lexeme-specific abbreviation arises, namely that all grapheme-specific abbreviations must be learned or consciously invented by a given scribe if they are to use them, whereas for lexeme-specific abbreviation the situation is more complicated. In the case of 'pure' lexeme-specific abbreviation, the logographic symbols in question must also be learned by a scribe before they can be used. In the case of 'impure' abbreviations, however, while more conventional examples such as ⟨ch'r⟩ may indeed have been learned by scribes either through training or by exposure to them, other instances such as the abbreviation of proper names by initials may have been more spontaneous. Knowledge that ⟨Gauvain⟩ can be abbreviated to ⟨G.⟩, for example, could be enough to allow a scribe to also abbreviate ⟨Lancelot⟩ to ⟨L.⟩, or ⟨Merlin⟩ to ⟨M.⟩, should those names appear sufficiently frequently within the text to ensure that a reader would be able to determine the intended reading from context.

Formal typologies such as that of Prou can therefore certainly be useful in the subcategorisation of our more functional typology, but it is the distinction between grapheme- and lexeme-specific abbreviation which should be our first-order typology when engaging with and analysing the medieval abbreviatory system. As we shall see in the following sections, the functional typology offers an interpretation of medieval abbreviation which is useful when we choose to take an analytical, rather than descriptive, approach.

2.3 Methodological considerations

2.3.1 Changing the object of analysis

Taking a functional approach to the typological classification of abbreviation in medieval French leads us to question what, in an analytical approach to such practices, should be the object of our analysis. Were we to base our approach on the typologies put forward by previous scholars, we may choose to look at the use of the abbreviatory markers themselves, as they appear within the manuscript. By contrast, the functional typology opens up the possibility of taking a different approach, looking instead at the lexemes and grapheme sequences that have been abbreviated, rather than the symbols used to do so.

An argument against the latter approach and in favour of the former is that abbreviated sequences cannot always be expanded with complete certainty, given the orthographic variation present in manuscripts written in medieval French, and the subtractive nature of the abbreviatory process. Both factors mean that the exact form taken by editorial expansions will always be influenced by subjective decisions made by the modern editor. Yvan G. Lepage, in his guide on editing medieval French texts, notes that when expanding abbreviations an editor ought to 'aligner la graphie [des mots abrégés] sur celle du copiste, qui écrit parfois ces mots au long dans son texte' (2001: 106). He illustrates this through the example of ⟨v^{9}⟩, which as we have seen above could be expanded by an editor as ⟨vos⟩, ⟨vous⟩, or ⟨vus⟩, the choice of which must be determined 'selon les habitudes du copiste et la couleur de sa *scripta*.'

There is certainly a degree of orthographic variation, as well as editorial subjectivity, involved in the preparation of any edition which contains expanded abbreviations, and this could negatively impact the quality of any analysis performed on it. For instance, if we were to analyse the abbreviation of ⟨os⟩, ⟨ous⟩, and ⟨us⟩ within a manuscript, treating them as wholly distinct sequences, then we might be led to conclusions which result less from varying uses of abbreviation than from orthographic variation, or even variation in the subjective preferences

of different editors involved in the preparation of a single edition. However, by simply treating these three forms as allographic variants of the same abbreviated unit, we are able to abstract away from such influences to gain an understanding of abbreviation usage which transcends surface-level orthographic and editorial variation. Thus, in the case of ⟨v^{9}⟩, we can create an abstract category containing the sequences ⟨os⟩, ⟨ous⟩, and ⟨us⟩, which allows us to look at the abbreviation of the word meaning 'you' without that analysis being impacted by external triggers of variation.

While external influence is a concern, then, it is far from insurmountable, and taking a slightly more abstract approach based on an understanding of allographic variation allows the benefits of this approach to come to the fore. An approach to the analysis of scribal abbreviation which focusses primarily on the sequences which have been abbreviated has the significant benefit of allowing a more granular examination of abbreviatory practices than an approach based on the abbreviatory markers themselves. For example, the so-called 'open a' ⟨ ᵘ ⟩[10] is an example of grapheme-specific abbreviation which can be placed atop another character in medieval French text to represent one of a number of possible abbreviated grapheme sequences, namely ⟨ua⟩ and ⟨ra⟩, such that there are therefore two possible grapheme sequences which are abbreviated by the same symbol.[11] An analysis based solely on the symbol used in the process of abbreviation could not account for the possibility that, for example, a given scribe may choose to abbreviate ⟨ua⟩ in this way, but not ⟨ra⟩. Instead, by looking at each of the sequences individually, we are able to gain a deeper understanding of trends within the data.

[10] This is also sometimes known as an 'astral r', a term coined by Richard Sharpe.
[11] The symbol also commonly appears as part of lexeme-specific abbreviations, further increasing the analytic granularity permitted by the functional approach.

Further, we should note that such examples are unlikely to be significantly impacted by editorial error resulting from confusion between the possible sequences, since the appropriate grapheme sequence can usually be determined beyond reasonable doubt by the surrounding graphemes. For example, an 'open a' on or around a ⟨q⟩ is overwhelmingly likely to be an abbreviation of the sequence ⟨ua⟩, as in the abbreviation of ⟨quant⟩, while the same symbol on or around a ⟨g⟩ is most likely to represent ⟨ra⟩, often to abbreviate ⟨grant⟩. Taking an approach which is focussed on the graphemes represented by particular abbreviations therefore allows us to delve deeper into data concerning the use of abbreviation without leading to a reduction in the validity of those data.

2.3.2 Methods of analysis

A functional approach based around what exactly is abbreviated, rather than the symbols used to do so, also allows us to perform statistical analyses of far greater validity than has previously been achieved. An example of analysis based on formal principles can be found in Geneviève Hasenohr's (2002: 84) statistical work on Old French abbreviation, which consists primarily of raw counts of abbreviatory symbols, measured per line of text or verse. Within it, she observes for instance that there are twelve abbreviations in the twenty-nine largely decasyllable lines of the ninth-century copy of the *Séquence de sainte Eulalie* found in Valenciennes, Bib. mun., 150. However, Hasenohr's approach is problematic for the simple reason that if we rely on how many abbreviations appear per line, or even per page, particularly for small datasets like the *Séquence*, we allow our data to be skewed by the text which is in that line, or on that page. The contexts in which a particular abbreviation may occur do not necessarily appear at the same rate from one line or page to the next, which could give the false impression that one section of text is intentionally more heavily abbreviated than another, when this is instead simply as a result of the fact that the text has more opportunities for certain types of abbreviation.

We must therefore take an approach to the analysis of medieval French abbreviation which considers observed abbreviations alongside the contexts in which such abbreviation could have appeared, but did not. By considering first and foremost the grapheme sequences elided during an instance of abbreviation, we are able to determine these contexts quite easily. For example, a sequence of a vowel followed by a nasal consonant can be counted as an instance of possible abbreviation, since it could have been abbreviated by the elision of the consonant and the insertion of a macron above the vowel.

Combining the total unabbreviated candidates with the observed number of actual abbreviations, we can calculate the total of how many opportunities a given section of text offers for abbreviation of a particular type. Dividing the number of observed abbreviations by this total gives the rate of abbreviation for that section of text, that is, the probability that, faced with a possible candidate for abbreviation, the scribe of that particular section will choose to abbreviate it. Use of this rate offers a far more interoperable and reproducible method than raw counts alone, allowing for comparison across texts of vastly different types, and giving greater certainty in its results. What matters in such a method is not whether a section of text includes multiple abbreviations for a given sequence, but rather whether, when such sequences do appear, the rate of abbreviation is higher or lower than that found elsewhere. Let us therefore apply this method to fr24432 as a case study, in order to display its utility.

2.4 Abbreviation in fr24432

As part of the project behind this thesis, and as discussed in the Introduction, I have produced a TEI P5-compliant digital edition of the whole of fr24432. This encoding proves an ideal dataset for a case study on the use of abbreviation in medieval French manuscripts, for several reasons. First, the edition includes the editorial expansion of abbreviations within the

manuscript within the <expan> and <ex> tags, which allows for digital extraction. Second, the fact that, as we saw in Chapter 1, fr24432 contains the work of some 11 different scribes means that it presents a significant opportunity for the examination of how abbreviation may vary between the work of different scribes, even when they are copying relatively comparable material. In addition to this, the length of fr24432, at 1775 columns of text, means that it provides a great many examples of both observed and possible abbreviation for our analysis.[12]

2.4.1 Macro-level analysis

A macro-level approach to fr24432 allows us to identify a number of interesting trends, and even relatively simple numerical approaches can be illuminating. For instance, an examination of the manuscript's 44,874 individual abbreviations reveals that fr24432 contains the abbreviation of 139 unique grapheme sequences. These vary significantly in frequency, with the most frequent abbreviation being that of nasal consonants, which are abbreviated on 14,798 occasions (33.0% of all abbreviated sequences), while some 113 sequences appear in fewer than 100 abbreviated instances.[13]

Looking first at the manuscript's examples of lexeme-specific abbreviation, we find that the most common such abbreviation is the abbreviation ⟨7⟩ for ⟨et⟩, found on 5,917 occasions (13.2%). Also particularly frequent are ⟨bn̄⟩ for ⟨bien⟩, itself occurring in 1,145 instances (2.6%), as well as ⟨m'lt⟩ for ⟨moult⟩, which is abbreviated 781 times (1.7%). Also appearing with notable frequency are ⟨n're⟩ and ⟨v're⟩ for ⟨nostre⟩ and ⟨vostre⟩ (423, 0.9%), ⟨n^{9}⟩ and ⟨v^{9}⟩ for ⟨nous⟩ and ⟨vous⟩ (398, 0.9%), ⟨.s'.⟩ for ⟨saint⟩ (249, 0.6%), ⟨ch'r⟩ for ⟨chevalier⟩ (158, 0.4%), and ⟨s't⟩ for ⟨sont⟩ (117, 0.3%). As already discussed, many of these forms such as ⟨.s'.⟩

[12] The manuscript contains 1,784 columns in total, but 9 of these are either blank or contain only contributions made by much later hands.
[13] For the purposes of this analysis, I have set numerals aside as a separate system, rather than as instances of abbreviation. It would be possible to argue, however, that the numeral system constitutes a subset of lexeme-specific abbreviation.

are reuses of abbreviations common in Latin text, while others such as ⟨ch'r⟩ appear to have been conceived independently, as forms uniquely for use in medieval French.

On some occasions, lexeme-specific abbreviations which are clearly limited to use in the vernacular nevertheless appear to have been created on the model of Latin abbreviations. ⟨C^{tre}⟩ for ⟨Chapitre⟩, found 39 times and only between ff. 317va and 326va, offers one such example. Within fr24432, we find 13 instances of ⟨C^{m}⟩ and similar variants to represent the Latin ⟨Capitulum⟩, which for their part all occur between ff. 320va and 325va. Given the proximity of this abbreviation to the instances of ⟨C^{tre}⟩, it appears that the scribe responsible for this section of text, or an earlier version text from which the copy in fr24432 was made, had two comparable lexeme-specific abbreviations available for use depending on whether they were copying in Latin or medieval French, reinforcing the suggestion that the medieval French abbreviatory system, while certainly developed from the Latin model, was nevertheless seen as distinct from it.

While there are a number of frequent lexeme-specific abbreviations within fr24432, grapheme-specific abbreviation appears on the whole to be the most frequent form of abbreviation used within fr24432. Besides the abbreviation of ⟨et⟩, grapheme-specific abbreviation accounts for the remainder of the 10 most frequent abbreviated sequences, namely the nasal consonants (14,798, 33.0%), ⟨ue⟩ (3513, 7.8%), ⟨ar⟩ (2803, 6.2%), ⟨ui⟩ (2630, 5.9%), ⟨er⟩ (1890, 4.2%), ⟨us⟩ (1787, 4.0%), ⟨ua⟩ (1432, 3.2%), ⟨com⟩ or ⟨con⟩ (1282, 2.9%), and ⟨ra⟩ (1254, 2.8%).

These higher rates of appearance of grapheme-specific abbreviation appear to result from the simple fact that the grapheme sequences which can be abbreviated by grapheme-specific abbreviation are far more frequent than the lexemes which can undergo lexeme-specific abbreviation. Notably, however, it appears to be a general truth that while grapheme-specific

abbreviation is more common overall, it is in fact lexeme-specific abbreviation which is the most frequent when we take into account the contexts in which it is possible.

For now, let us take two relatively simple examples to illustrate this, namely the grapheme-specific abbreviation of nasal consonants, and the lexeme-specific abbreviation of ⟨et⟩, which are the two most common abbreviations within fr24432. By calculating the rates of abbreviation for both these sequences,[14] we find that while there are more than twice the number of examples of the abbreviation of nasal consonants compared to that of ⟨et⟩, the rate of abbreviation of the latter sequence is in fact much higher than that of the former, with nasal consonants abbreviated at a rate of 19.6%, while ⟨et⟩ is abbreviated at a rate of 41.8%. In less frequent cases of lexeme-specific abbreviation the difference is even more stark, with not a single unabbreviated instance of ⟨Chapitre⟩ occurring between ff. 317va and 326va.

These macro-level findings therefore reinforce the utility of our distinction between lexeme- and grapheme-specific abbreviation, in that the data suggest that scribes treated these two types of abbreviation differently. Lexeme-specific abbreviation in fr24432 appears with high rates of abbreviation, but only in the contexts where it is possible. The limited contexts in which such abbreviation can appear mean that on the whole, and with the exception of ⟨7⟩ for ⟨et⟩, lexeme-specific abbreviation appears less frequently than grapheme-specific abbreviation, even though the rates of abbreviation in the latter type are often much lower.

These data therefore allow us to take issue with the claims which have been made by a number of scholars, including Hasenohr (2002: 82–83), that abbreviation in medieval French is far less frequent than that found in Latin. Our data here are insufficient to allow us to refute this claim in its entirety, but they do allow us to nuance it slightly by suggesting that

[14] For details of what is considered a possible candidate for abbreviation for each of these sequences, see the discussion below.

medieval French texts use less of the lexeme-specific abbreviation which is common Latin texts, given the absence of the broad and highly-developed system present for Latin, while grapheme-specific abbreviation remains widespread. In doing so, we may nuance Mazziotta's (2008: §67) claim that 'les abréviations qu'on rencontre dans les manuscrits français sont généralement d'un niveau de complexité inférieur à celui de celles qu'on rencontre dans les manuscrits latins,' noting that this reduction in complexity stems from the lesser emphasis placed on lexeme-specific abbreviation in medieval French text when compared to Latin.

Of course, an in-depth comparative study of Latin and medieval French abbreviation would be necessary to confirm this, but at the very least our division between lexeme- and grapheme-specific abbreviation proves helpful in generating hypotheses about the existence of a distinction between the practices associated with each language.

2.4.2 Micro-level analysis

So far, we have looked at fr24432 in its entirety. As we have also seen through the abbreviation of ⟨Chapitre⟩, though, the separate sections of text contained within the codex are not entirely homogeneous. Instead, they vary in form and content, and have been copied by a number of different individuals. There is therefore also a value in taking a micro-level approach, looking at how the use of abbreviation changes across the manuscript, and what factors may explain any variation.

Discrete variables

Much of the variation which occurs within manuscript texts can be explained by variation in the preferences of the individual scribes who copied it. Orthography, for instance, varies based on the training and origins of the individual copyist, and the period in which they were writing. Even scribes who hail from similar areas at similar points in history, and who appear to have received comparable training, vary in their spelling apparently on the basis of

personal preference. It is this variation which has led Gabriella Parussa and Richard Trachsler (2006: 190) to argue for the possibility of identifying a 'profil orthographique,' an *orthographic profile*, for a given scribe based on how they habitually spell certain word forms.

Drawing on this approach, I have argued that it is also possible to identify *abbreviatory profiles* for individual scribes, even those working closely together on the same text. That is, it is possible to identify a scribe's 'preferred methods of abbreviating certain graphemes and lexemes, and the frequency with which he chooses to do so' (Dows-Miller 2022: 61–62).

By separating the text of fr24432 into the 20 scribal stints of fr24432 which we identified in Chapter 1, we are able to identify such abbreviatory profiles by considering evidence of discrete variation, that is, examples of distinct methods of abbreviation used by the manuscript's scribes for the same sequences of graphemes.

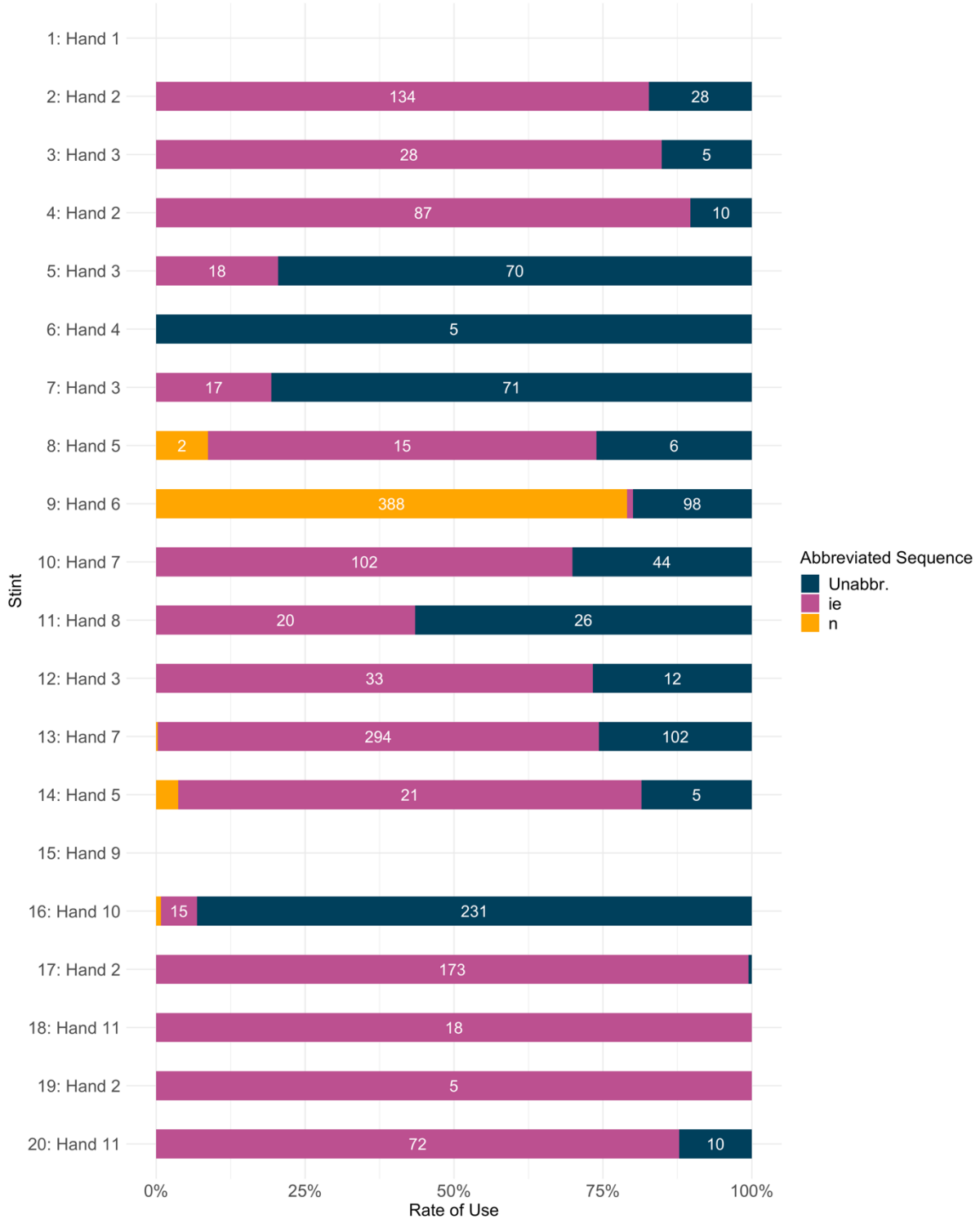
⟨bien⟩

Perhaps the clearest such example anywhere in the manuscript comes from the abbreviation of the sequence ⟨bien⟩, used as an adverb to mean 'well' and a singular noun to mean 'goodness', and appearing as a plural noun ⟨biens⟩ to mean 'goods'. Across the manuscript, when ⟨bien(s)⟩ is abbreviated, we find two possible methods of abbreviation, namely the abbreviation of ⟨ie⟩ through the use of a macron bar atop ⟨n⟩, as in ⟨bn̄⟩, and the abbreviation of ⟨n⟩ through a macron bar above the second vowel, as in ⟨biē⟩. The former method is a reapplication of the common Latin abbreviation for ⟨bene⟩ (Cappelli 1982: 9), and is an example of lexeme-specific abbreviation, while the latter is simply an application of the common grapheme-specific abbreviation of a nasal consonant using a macron bar.

Figure 2.31 lays out the usage of each method across the 20 scribal stints of fr24432, and allows us to identify the variation in the preferred abbreviated sequence for each stint. Where

a stint contains no data, this is because there are no abbreviated or unabbreviated examples of the sequence in question, which usually results from the stint being only a few folios in length.

Figure 2.31 – use of different abbreviatory methods in the abbreviation of ⟨bien⟩ *and its variants.*



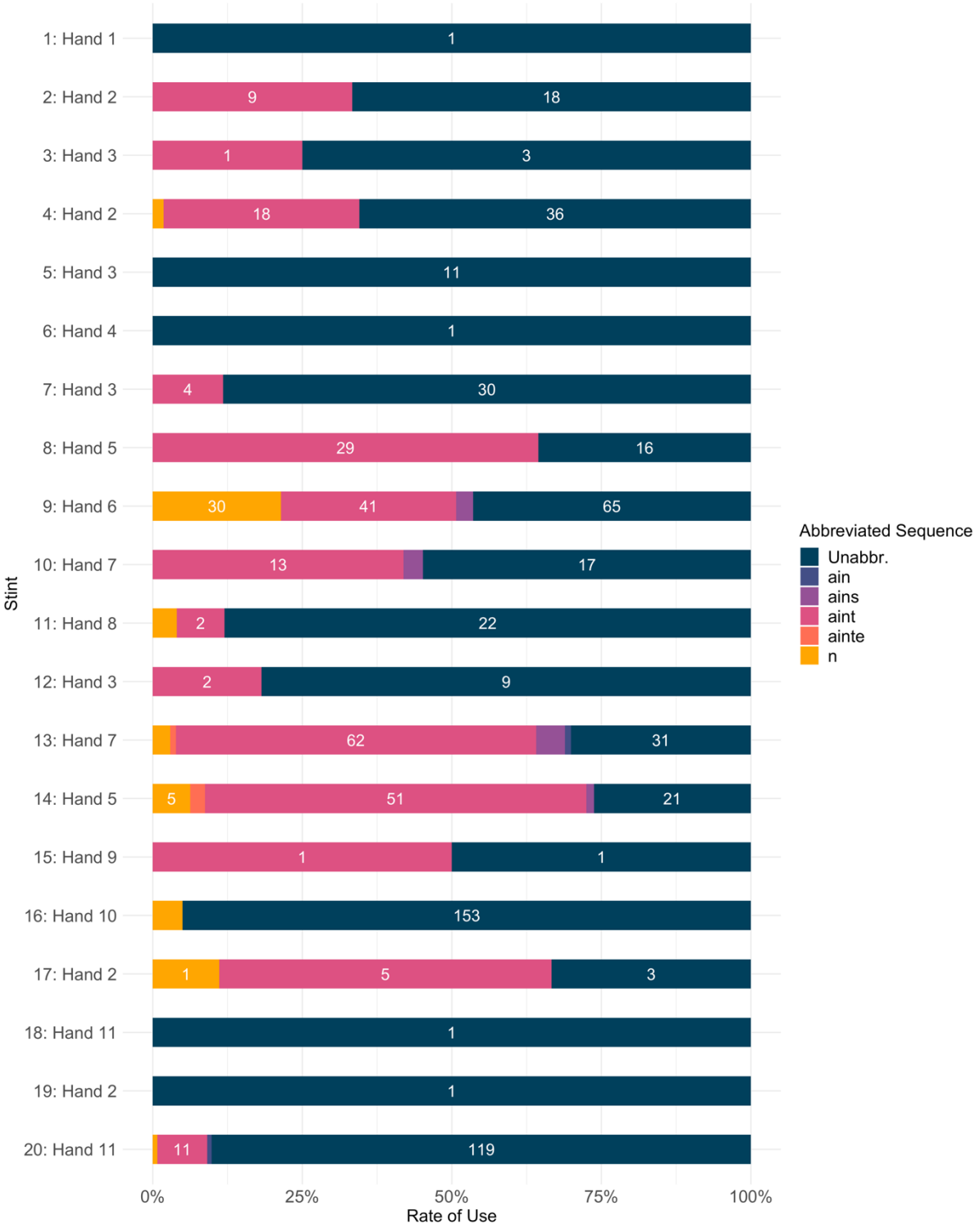
Examining the data, we find that the overall trend within fr24432 is to abbreviate ⟨bien(s)⟩ at a very high rate, with only hand 4, hand 3 (in their two middle stints: stints 5 and 7), hand 8, and hand 10 preferring to retain the sequence in an unabbreviated form. What is perhaps most notable, however, is the fact that the scribes of fr24432 are almost wholly consistent in which method of abbreviation they choose. Almost without exception, the preferred method of abbreviation within the manuscript is to use the lexeme-specific ⟨bn̄⟩, while only hand 6 prefers the more generic, grapheme-specific ⟨biē⟩. This reveals the significant impact that scribal preference can play on the choice of abbreviatory methods in fr24432, given that the texts copied by hand 6 are broadly comparable to those in adjacent stints, and may even have come from the same exemplars.

While the abbreviatory practice observed within the work of hand 6 is therefore clearly the result of the scribe's individual preference, the influence of abbreviations present in exemplar copies is made clear by the fact that a number of scribes, including hand 6, are not wholly consistent in the method of abbreviation they choose. While hand 6 contains 5 abbreviated forms using ⟨bn̄⟩, we also find the grapheme-specific method in the work of hand 5 (2 in stint 8, 1 in stint 14), hand 7 (1 in stint 13), despite both hands generally preferring the lexeme-specific method. Two further examples of the grapheme-specific method are also found in the work of hand 10 alongside a greater number of lexeme-specific examples, and yet the scribe here appears to prefer not to abbreviate ⟨bien⟩ at all. Hand 3, meanwhile, varies significantly across their 4 stints (3, 5, 7, and 12) in the rates at which they leave ⟨bien⟩ unabbreviated, with stints 3 and 12 containing much higher rates of abbreviation than stints 5 and 7, which could also result from variation in the rates of abbreviation present in the exemplars. If we take this underlying variation to be an indication of preferences present in the scribes' exemplar manuscripts, we are therefore able to suggest that while scribes' abbreviatory profiles and practices were predominantly determined by their individual

preferences, they could on occasion be influenced by the practices present in the texts from which they copied.

⟨saint⟩

Figure 2.32 – use of different abbreviatory methods in the abbreviation of ⟨saint⟩ *and its variants.*



A similar division between preferences for lexeme- and grapheme-specific methods of abbreviation can be found in the case of ⟨saint⟩ and its inflectional variants, meaning 'saint' as a noun or 'holy' as an adjective. A number of different sequences are abbreviated from within this word, but they boil down to two distinct methods. The first, a lexeme-specific abbreviation, involves what in Prou's typology would be termed abbreviation through the use either of initials or of contraction, with ⟨saint⟩ reduced to a single ⟨s⟩, often capitalised, sometimes with the addition of an inflectional ⟨e⟩, and usually with surrounding *puncti* and a following superscript mark like a modern apostrophe, to give ⟨.s'.⟩ or ⟨.s'e.⟩.[15] As with ⟨bn̄⟩, this method is one drawn from Latin abbreviation, and appears to be the primary preferred method across much of fr24432, as displayed in Figure 2.32.

When used in a masculine plural form, the unabbreviated form of the word is usually ⟨sains⟩, and less frequently ⟨sainz⟩. When abbreviated this number inflection is unmarked, as in the work of hand 6 (4 instances), hand 7 (1 in stint 10, 5 in stint 13), and hand 5 (1 in stint 14). By contrast, in the feminine form, the inflectional ⟨e⟩ usually appears unabbreviated after the initial ⟨s⟩, with the exception of a few instances in the work of hand 7 (1 in stint 13) and hand 5 (2 in stint 14) where, assuming this is not simply a case of scribal omission, the inflectional ⟨e⟩ is apparently included within the abbreviated graphemic structure rather than outside it. In one instance in the work of hand 7 (stint 13) and hand 11 (20), the final ⟨t⟩ is left unabbreviated, in both cases in the presence of an unabbreviated inflectional ⟨e⟩, giving the form ⟨.s'te.⟩.

The alternative method for the abbreviation of ⟨saint⟩ and its variants is again through the application of a grapheme-specific rule, through which the ⟨n⟩ is replaced by a macron bar

[15] We note another benefit of the functional typology over the formal one here, namely that the addition of the ⟨e⟩ causes the abbreviation to arbitrarily change category in the formal typology but not in the functional one.

atop the preceding vowel. This is far from the most frequently used method across fr24432 for this particular sequence, with all but hand 10 preferring the lexeme-specific model, and yet it is nevertheless still interesting that the rate of use of this method within the work of hand 6 is by far the greatest in the codex, given that it was this hand that preferred the same approach for ⟨bien(s)⟩. It therefore appears plausible that hand 6 had a greater predilection for grapheme- over lexeme-specific abbreviation than their colleagues, even if they did still occasionally adopt more standard, lexeme-specific methods of abbreviation, perhaps under the influence of forms present in their exemplars.

⟨sont⟩

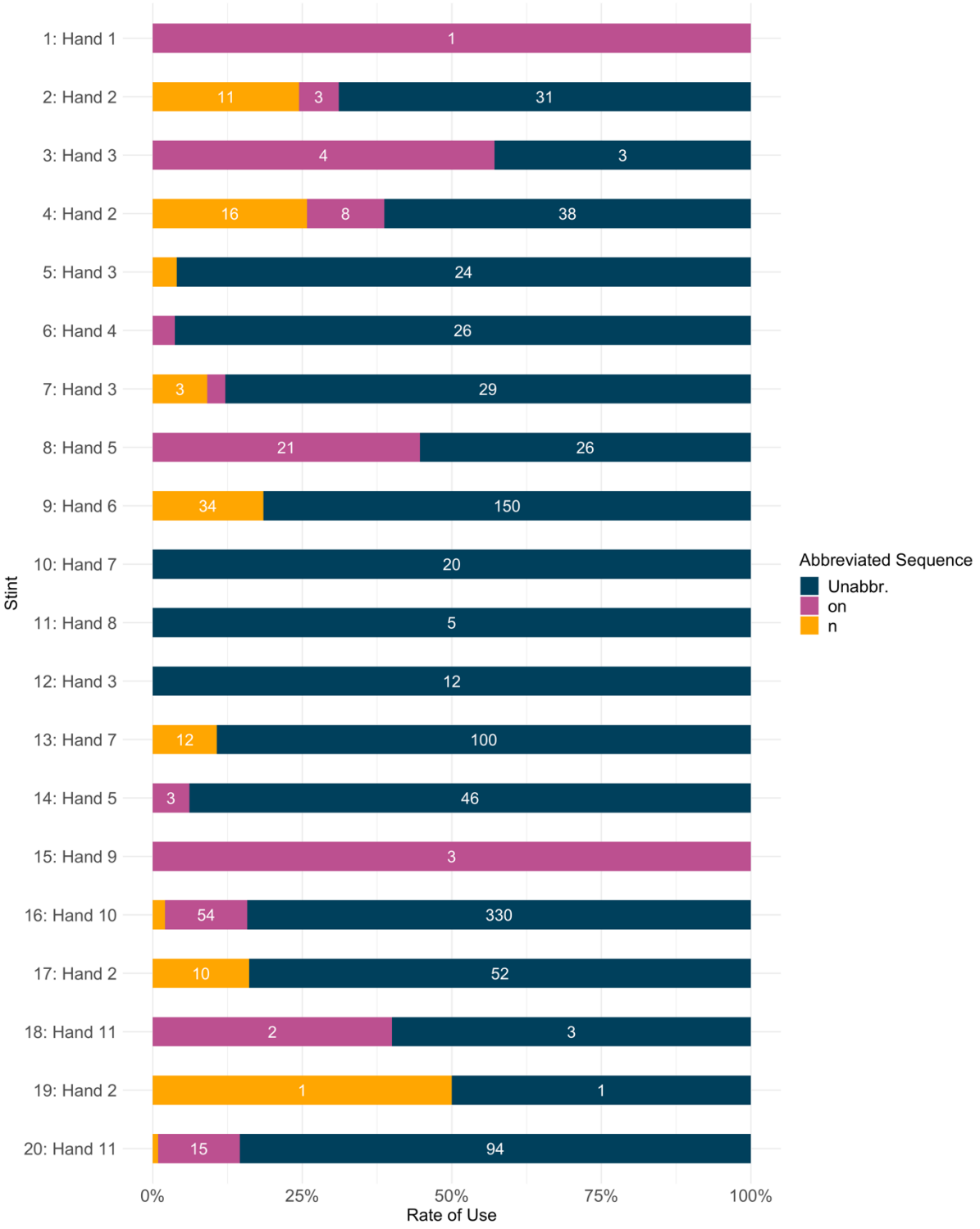
The single inflected form ⟨sont⟩ can again be abbreviated by one of two methods, one of which is grapheme-specific, and the other of which is a lexeme-specific abbreviation of the 'impure' type. The grapheme-specific option involves, once again, the use of a macron bar to indicate the abbreviation of ⟨n⟩, while the lexeme-specific method involves the use of ⟨s't⟩, an abbreviation commonly found in Latin for the equivalent ⟨sunt⟩ (Cappelli 1982: 13). Looking at the rates of usage in Figure 2.33, we find that the preference by far is to leave ⟨sont⟩ unabbreviated, but when it is abbreviated scribes tend to display a preference in one direction or another, with hands 2 and 6 preferring the lexeme-specific approach,[16] and hands 5 and 11 preferring the grapheme-specific approach.

Just as we observed a difference in the rates of abbreviation of ⟨bien⟩ across the stints of hand 3, here we also notice a distinction in the method of abbreviation which the scribe prefers. In their first stint (stint 3), hand 3 displays a strong preference for the lexeme-specific method, while in the remaining stints (stints 5, 7, and 12), there is very little abbreviation of ⟨sont⟩ whatsoever. Combining this with the differential rates for ⟨bien⟩, we may hypothesise that the

[16] The same may also be said of hands 1 and 9, but the sample sizes are too small to do so with any certainty.

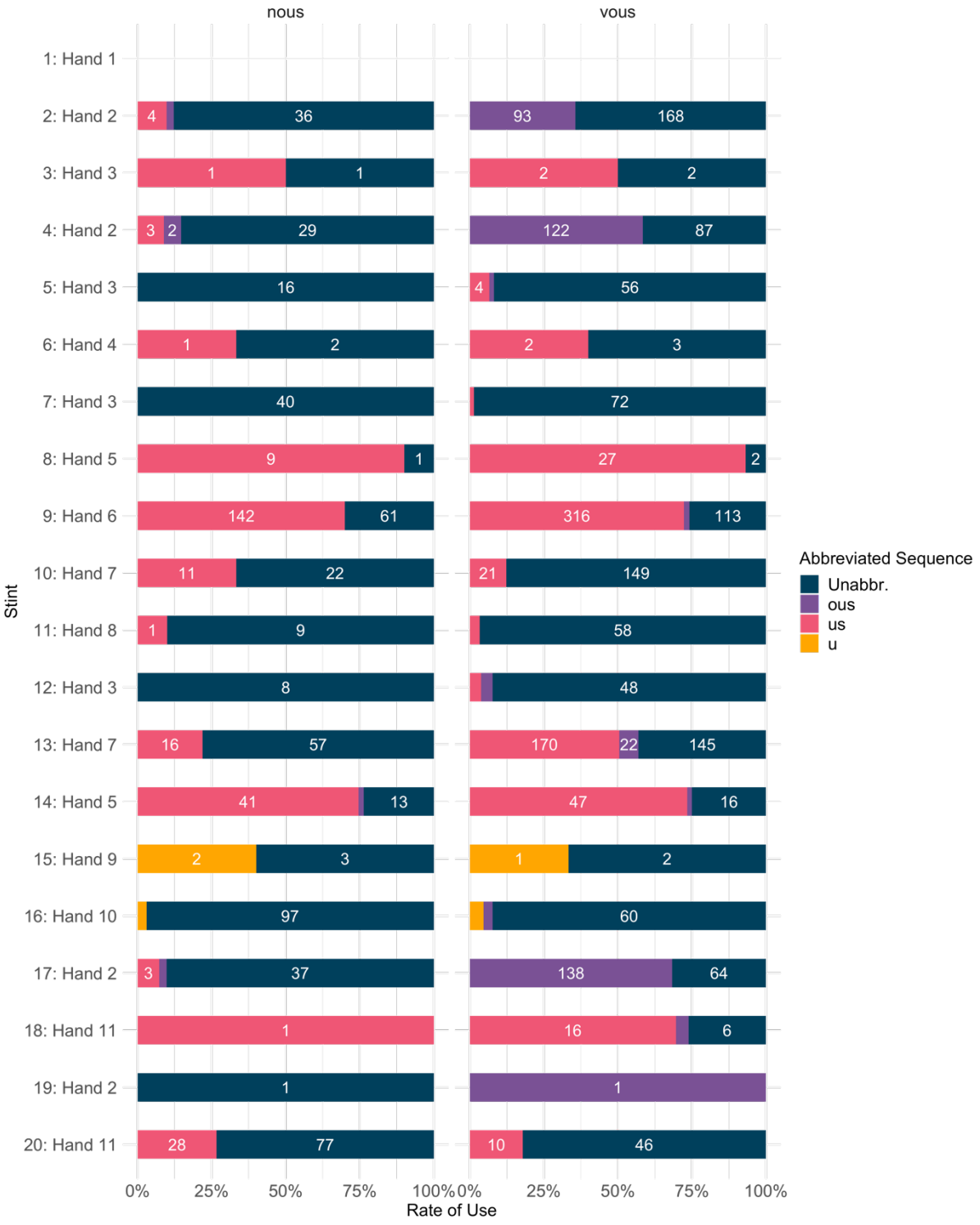
exemplar(s) from which hand 3 copied were either 3 distinct manuscripts with different preferences (with stints 5 and 7 from the same exemplar), or at the very least that the single exemplar from which they copied their stints contained 3 distinct abbreviatory practices.

Figure 2.33 – use of different abbreviatory methods in the abbreviation of ⟨sont⟩.



⟨nous⟩ and ⟨vous⟩

Figure 2.34 – use of different abbreviatory methods in the abbreviation of ⟨nous⟩ *and* ⟨vous⟩.



The pronouns ⟨nous⟩ and ⟨vous⟩ are abbreviated in fr24432 using one of two possible methods, both of which are examples of grapheme-specific abbreviation. The first method

involves the use of the symbol ⟨⁹⟩, which is taken to stand in some cases for ⟨ous⟩, in forms like ⟨n⁹⟩ and in others for ⟨us⟩, as in ⟨no⁹⟩.[17] The latter use of ⟨⁹⟩ is the standard graphemic referent for this marker (Cappelli 1982: 13), but this can be modified in certain lexical contexts, given the fact that ⟨n⁹⟩ can be used to represent the Latin ⟨nos⟩ (Cappelli 1982: 15), itself the equivalent of the medieval French ⟨nous⟩. The overwhelming preference in fr24432 for the ⟨no⁹⟩ type, as observed in Figure 2.34, suggests however that this modified use was non-standard, and thus that ⟨⁹⟩ was seen by most of the scribes of the codex as a marker with fixed graphemic reference.

It is therefore notable that, while there is some variation in usage between these two variants of the same method, hand 2 displays a consistent preference for the abbreviation of ⟨ous⟩. Significantly, however, this is only the case for ⟨vous⟩, which hand 2 never abbreviates as ⟨vo⁹⟩. The pronoun ⟨nous⟩, meanwhile, is most commonly left unabbreviated, and the abbreviation of ⟨us⟩ outweighs that of ⟨ous⟩ when abbreviation does occur. From this, we are able to reach a general understanding that it is wholly possible for scribes to have different preferences for the abbreviation of such comparable words, as well as a more specific sense that hand 2, by following the abbreviatory practice observed for some Latin pronouns, may well have had a greater awareness of, or affinity with, the norms of Latin abbreviation than their colleagues, even though the texts being copied are in French.[18]

Hands 9 and 10 display an intriguing alternative method for the abbreviation of ⟨nous⟩ and ⟨vous⟩, through the use of a macron bar which, rather than representing a nasal consonant, as is standard, instead represents the character ⟨u⟩. There are relatively few such examples, with

[17] On 14 occasions across the manuscript, ⟨n⁹⟩ also appears as an abbreviated form of ⟨nus⟩, usually in a negated structure meaning 'none'.
[18] The similarity with the standard forms ⟨nus⟩ and ⟨vus⟩ in Anglo-Norman could also reveal some scribal familiarity with Insular copying traditions, although there is little else in the scripta of hand 2 to support this.

3 appearing across the work of hand 9, and 6 in the work of hand 10, but the fact that this method is shared by both scribes does reinforce the sense, as discussed in Chapter 1, that the work of hands 9 and 10 is a later addition to fr24432, with the two hands showing comparable scribal practices which are distinct from those found in the rest of the codex.

⟨com⟩ and ⟨con⟩

Turning from the abbreviation of full words to that of partial grapheme sequences, we observe a distinction in the methods used to abbreviate the sequences ⟨com⟩ and ⟨con⟩. As observed for ⟨bien⟩, ⟨saint⟩, and ⟨sont⟩, one method involves the generic use of a macron bar atop the ⟨o⟩ to represent a nasal consonant. The other, meanwhile, involves the use of a special character ⟨9⟩, which represents the full sequence ⟨com⟩ or ⟨con⟩.[19]

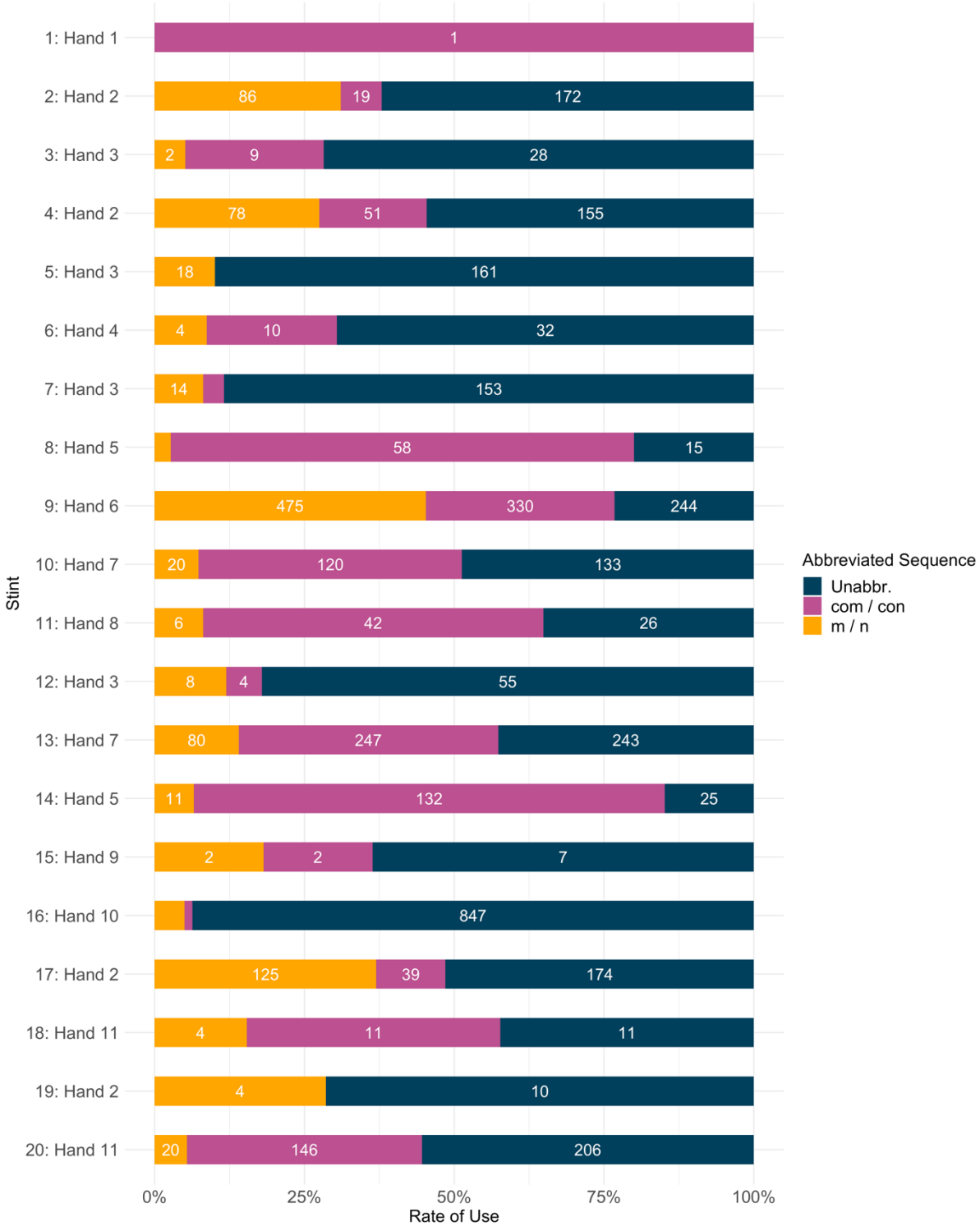
From Figure 2.35, we see that both methods appear in alternation for almost every scribe, with the exception of individual stints of hands 3 (stint 5) and 2 (stint 19), despite the presence of variation in other stints by the same hands.[20] Hand 3 is once again notable in that the scribe varies in their preferred method of abbreviation across their four stints, with stints 5, 7, and 12 displaying a preference for the abbreviation of the nasal consonant, while in stint 3 they show a clear preference for the abbreviation of ⟨com⟩ or ⟨con⟩ in their entirety, adding further weight to the suggestion that hand 3's stints were copied from either more than one exemplar manuscript, or a single exemplar displaying a variety of abbreviatory methods.

In spite of this variation, we observe that on the whole, most scribes appear to have a preference for one method or the other. Hands 5, 7, 8, and 11 display a strong preference for abbreviation of the full sequence, while hands 2 and 6 prefer simply to abbreviate the nasal

[19] ⟨com⟩ and ⟨con⟩ are treated here as two allographic variants of the same sequence, given that the decision to expand an abbreviation as one or the other is ultimately editorial.
[20] Hand 1 must be ignored here, since the sample size (of 1) is too small to draw any real conclusions.

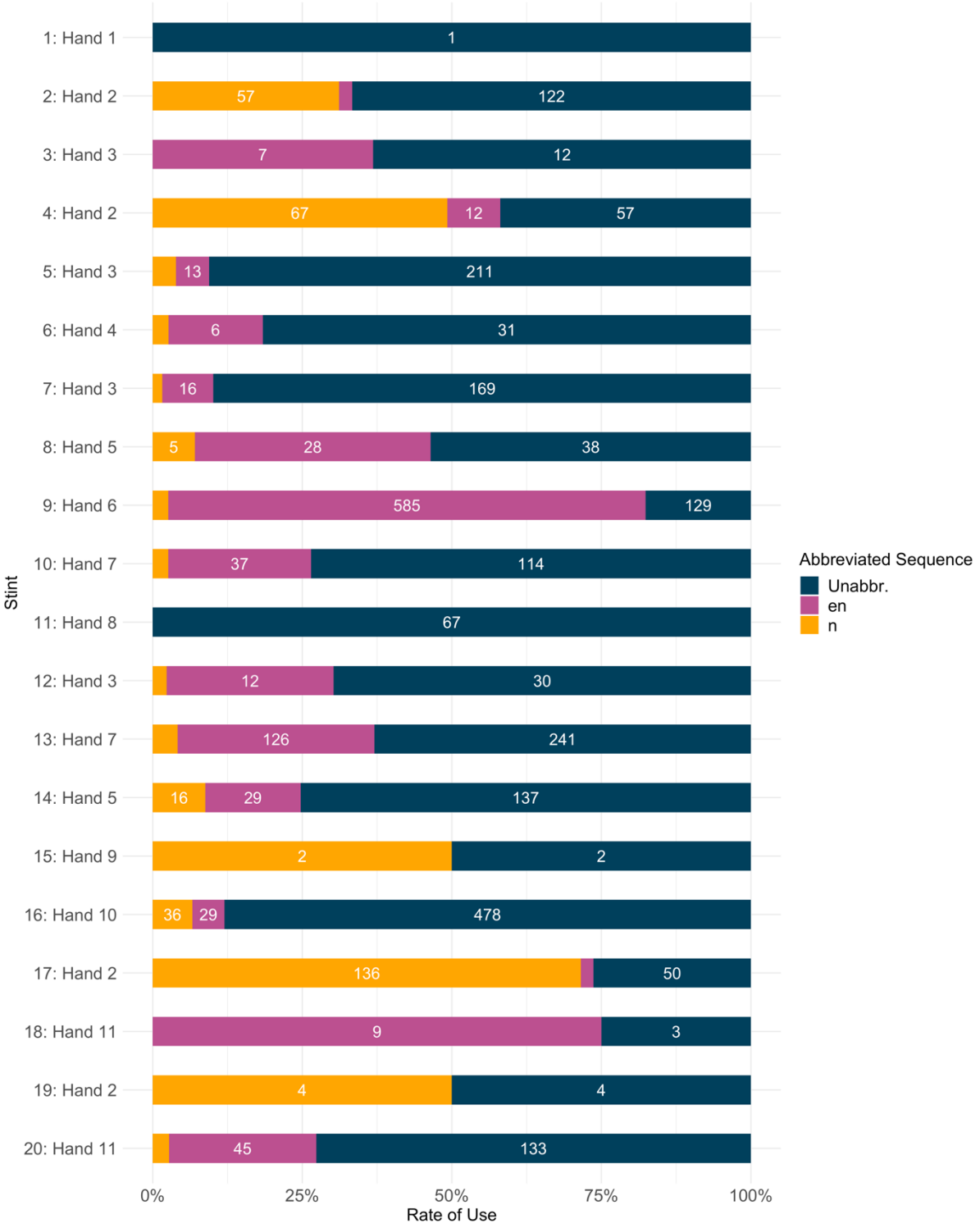
consonant. Meanwhile, as we have already seen elsewhere, hand 10 prefers not to abbreviate very many sequences at all.

Figure 2.35 – use of different abbreviatory methods in the abbreviation of ⟨com⟩ *and* ⟨con⟩.



Forms ending in ⟨ment⟩

Figure 2.36 – use of different abbreviatory methods in the abbreviation of word-final ⟨ment⟩ *and its variants.*



Variation in the abbreviation of grapheme sequences that form part of a word can also be seen in the abbreviation of word-final ⟨ment⟩, and in its inflected variants ⟨mens⟩ and

⟨menz⟩.[21] Both methods of abbreviation involve use of the macron bar, with the more generic approach involving the abbreviation of the ⟨n⟩, as in ⟨mēt⟩, ⟨mēs⟩, and ⟨mēz⟩, while a more idiosyncratic usage involves the abbreviation of ⟨en⟩, as in ⟨m̄t⟩, ⟨m̄s⟩, and ⟨m̄z⟩.

Ignoring the small dataset for hand 9, hand 2 is the only scribe to show a significant preference for the use of the macron bar specifically to represent ⟨n⟩, which lines up with their similar preference for the use of the more generic method in the abbreviation of ⟨com⟩ or ⟨con⟩. The abbreviation of ⟨en⟩, meanwhile, is much more common than that of ⟨n⟩ for the abbreviation of ⟨ment⟩ and its variants in the work of hands 3, 4, 5, 6, 7, and 11. This abbreviation of ⟨en⟩ therefore overlaps with the manuscript's 5 instances of the use of ⟨m̄⟩ as an abbreviation of ⟨mem⟩ before a bilabial consonant, as in ⟨rem̄brance⟩ for ⟨remembrance⟩, which is used 4 times within the work of hand 3 (2 in stint 3, 2 in stint 12), and once in the work of hand 11 (stint 20). Such an application of the macron bar beyond instances of ⟨ment⟩ to represent ⟨e⟩ plus a following nasal consonant therefore reinforces the sense that this feature is a result of scribal preference rather than a residual result of the exemplar from which they copied.

The greatest rates of abbreviation, however, come in the work of scribe 6, who once again reveals their preference for frequent use of abbreviation. Interestingly enough, on this occasion, the scribe's preferred method is not the most generic available, unlike that which we observed for ⟨bien⟩, such that it becomes even more plausible that the motivation behind the choice of hand 6 in that case was based on a preference for grapheme-specific forms.

[21] Variant spellings with ⟨a⟩ instead of ⟨e⟩ have also been included here.

Forms ending in ⟨ion⟩

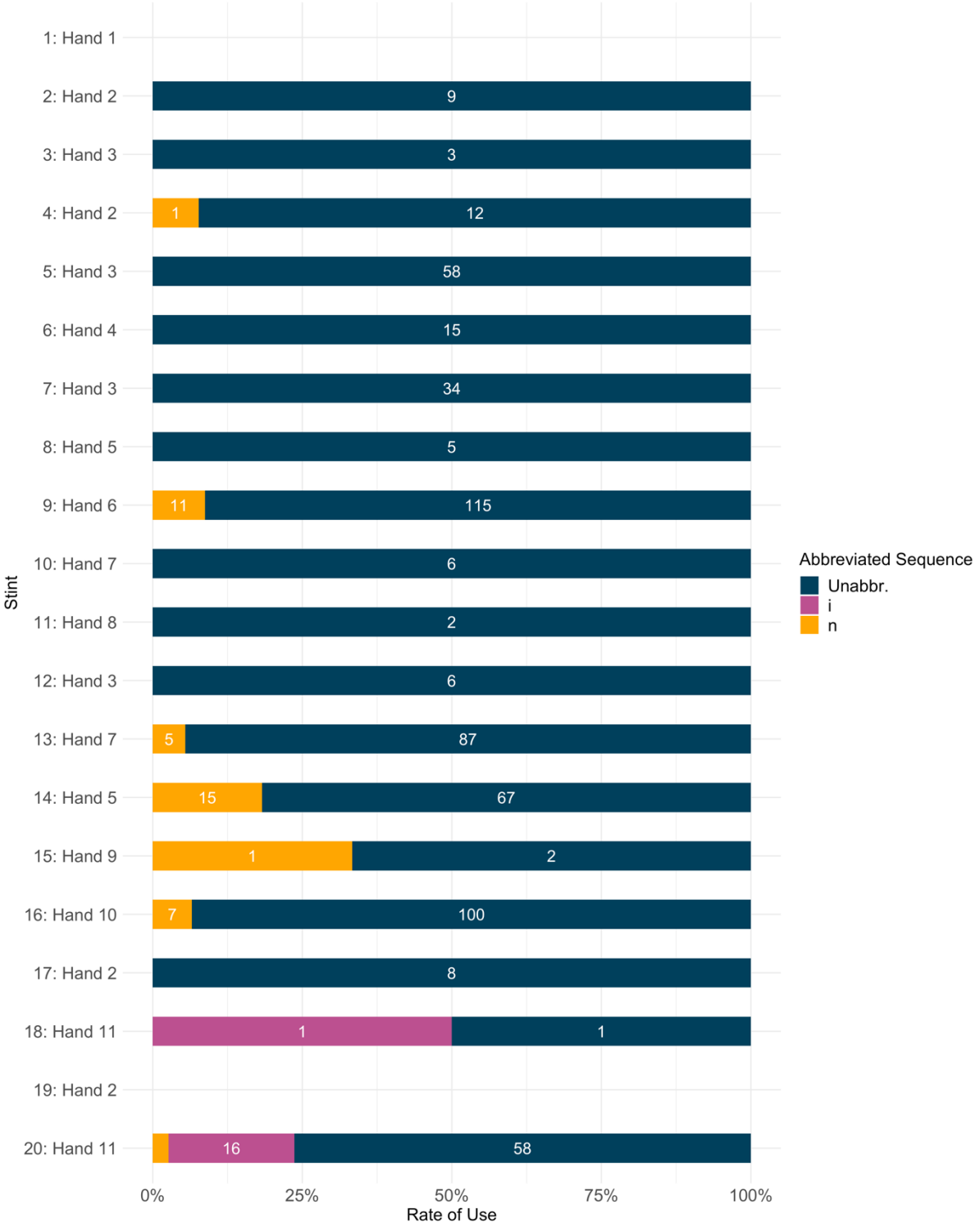
Our final sequence of note is again word-final, specifically ⟨ion⟩, which may be followed by an inflecting ⟨s⟩ or ⟨z⟩.[22] What is most notable about such sequences, as presented in Figure 2.37, is that they are largely entirely unabbreviated, with 10 of the 18 stints in which abbreviation would be possible containing no abbreviated examples. Perhaps unsurprisingly given the methods of abbreviation discussed so far, where abbreviation does occur, it is usually through the use of a macron to represent ⟨n⟩, as in ⟨tiō⟩, with such forms appearing in stints by hands 2 (stint 4), 6, 7 (13), 5 (14), 9 (15), 10, and 11 (20).

A rare and intriguing alternative method of abbreviation within the codex is that found in both stints of hand 11, in which a macron bar is used above an ⟨o⟩, not to represent a following nasal consonant, but instead a preceding ⟨i⟩, as in ⟨tōn⟩. This method could arguably be the result of the abbreviatory preferences of hand 11, but it is also plausible that these abbreviations could result from those present in a single exemplar.

The latter possibility is supported by the fact that this abbreviation, while rare in fr24432, is attested by Cappelli (1982: 13) as a method by which the Latin word-final sequence ⟨tio⟩ may be abbreviated. Items 89 and 90,[23] found within the final stint of hand 11, contain a number of Latin citations (accompanied in all cases by a medieval French translation), and therefore these appear to be the work of a more Latinate copying tradition than much of the rest of the manuscript, where Latin is much less common. These unusual abbreviations, then, could well have been present in the chain of transmission from soon after the texts' composition.

[22] In order to avoid including an inordinate number of first-person plural verb forms, which do not appear to abbreviate according to the same rules, for this analysis I have chosen only those words which end in ⟨cion⟩, ⟨sion⟩, ⟨tion⟩, and ⟨xion⟩, as well as variants of those forms ending in ⟨s⟩ or ⟨z⟩.
[23] 89: *Les proverbes de Salomon*, a prose presentation of proverbs and sayings; 90: An anonymous prose translation of *De misera humanae conditionis* by Cardinal Lotario dei Segni, the future Pope Innocent III.

Figure 2.37 – use of different abbreviatory methods in the abbreviation of word-final ⟨cion⟩, ⟨sion⟩, ⟨tion⟩, ⟨xion⟩, *and their variants.*



If hand 11 was indeed copying abbreviated forms already present in their exemplar, then that would suggest that item 88 was also present within the same exemplar, since it too contains an example of the Latinate abbreviatory method. This would be a surprise, given that item

88, *Le miroir de vie et de mort*, as a verse dream allegory, is significantly different from the prose texts that follow it, thus revealing the significance of an analysis based on abbreviatory practices as a means of determining the possible means by which a manuscript may have been compiled.

Continuous variables

Having considered the abbreviation in fr24432 through a discrete lens, we may also take a more continuous approach to the manuscript's abbreviation, looking at the rates of the abbreviation of a given sequence change across the whole codex.

In order to take such a continuous approach, we must first divide a given manuscript into units which will contain a sufficient number of both actual abbreviated examples and possible contexts for abbreviation. In most instances an individual leaf is sufficient. By taking such a narrow approach, however, we necessarily narrow the scope for which abbreviations can be considered, since for some grapheme-specific abbreviations, and particularly for lexeme-specific abbreviation, there simply will not be enough possible candidates for abbreviation, whether that abbreviation occurs or not, to allow for statistically valid analysis at such a fine-grain level. It is therefore necessary to limit our focus to only the most widespread abbreviations. In this case, I have chosen to exclude all abbreviated sequences which form less than 1% of the total observed abbreviations, since this allows us to focus only on the abbreviations which are likely both to have possible candidates for abbreviation, as well as actual abbreviated examples, on most leaves of the codex. This narrows the list of abbreviated sequences from 139 to 14, namely (in decreasing order of frequency) the nasal consonants, ⟨et⟩, ⟨ue⟩, ⟨ar⟩, ⟨ui⟩, ⟨er⟩, ⟨us⟩, ⟨ua⟩, ⟨com⟩ / ⟨con⟩, ⟨ra⟩, ⟨ie⟩, ⟨en⟩, ⟨re⟩, and ⟨ou⟩.

One further complicating factor comes in determining what counts as a possible candidate for abbreviation. While there are standard contexts in which many of the abbreviated sequences

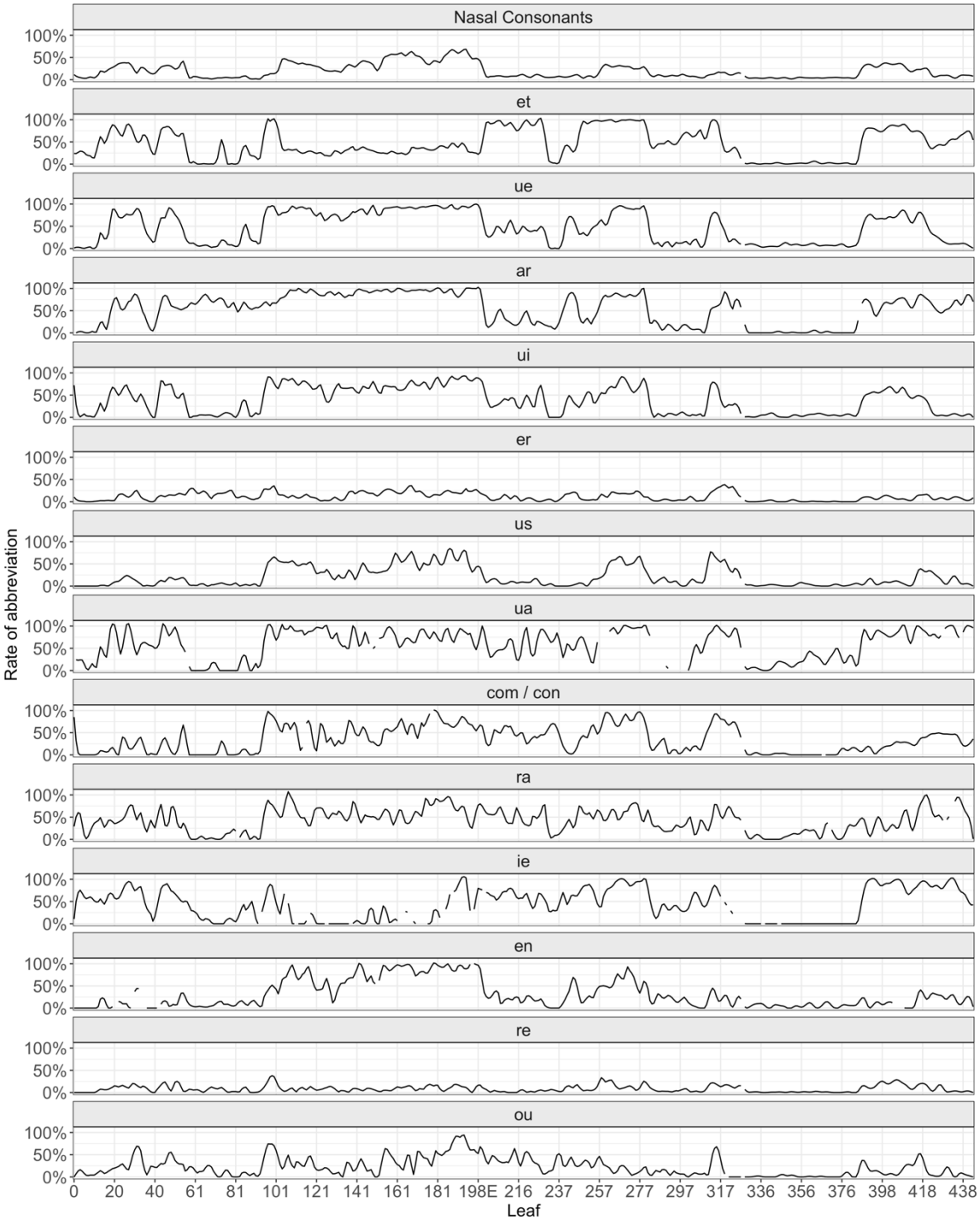
are possible, such as the fact that ⟨et⟩ is usually abbreviated as a full word, and that ⟨ui⟩ is most commonly abbreviated after ⟨q⟩, even standard abbreviations occasionally appear in non-standard uses, such as the rare cases in which ⟨7⟩ is used as a grapheme-specific marker for ⟨et⟩ as part of a larger graphemic structure. If, in determining possible candidates for abbreviation, we were simply to extract all the contexts in which abbreviation is observed, and assume that abbreviation would have been possible any time these contexts appeared, then the rare, non-standard uses of certain abbreviations would skew the data by including candidates for abbreviation which most scribes would never have considered possible.

For example, given the presence of a macron-type mark atop an ⟨m⟩ found on f. 224va to represent a following nasal consonant, if one were to assume that every nasal consonant following an ⟨m⟩ or an ⟨n⟩ is a possible candidate for abbreviation through that method, then one would add several thousand possible candidates across fr24432. This would be a mistake, however, given that there is only a single instance of this use of the macron in the whole manuscript, and so this is clearly a non-standard use of a standard abbreviatory method.

Determining which observed contexts of abbreviation to include and which to ignore is therefore important, but given the thousands of abbreviations which occur in even a single manuscript, doing so manually is unfeasible. Instead, we can again set a threshold for which contexts to exclude and which to include, which has been set here at 1%. An observed context in which a given abbreviation appears is therefore only considered valid as a possible context for abbreviation anywhere in the manuscript if at least 1% of the observed abbreviations appear in the same context. For example, in fr24432 there are 7 different characters which appear immediately before the abbreviated grapheme string ⟨ar⟩, and 20 which appear afterwards. By applying the 1% threshold, we are able to narrow this list down to just one character before (⟨p⟩) and 10 after (⟨a⟩, ⟨d⟩, ⟨e⟩, ⟨f⟩, ⟨l⟩, ⟨o⟩, ⟨s⟩, ⟨t⟩, ⟨u⟩, and ⟨ ⟩—

that is, the end of the word). Through this refinement, the data are less likely to be skewed by allowing rare contexts to define the candidates for abbreviation.

Figure 2.38 – rates of abbreviation for the fourteen most frequent abbreviated sequences per leaf in fr24432.



Once we have identified the abbreviated sequences of interest, and the contexts of abbreviation, it is possible to calculate the rates at which a given sequence is abbreviated

across all leaves of the manuscript. These data are presented in Figure 2.38 for our 14 most frequently abbreviated sequences in fr24432 and allow us to observe significant variation in their use across the codex.[24]

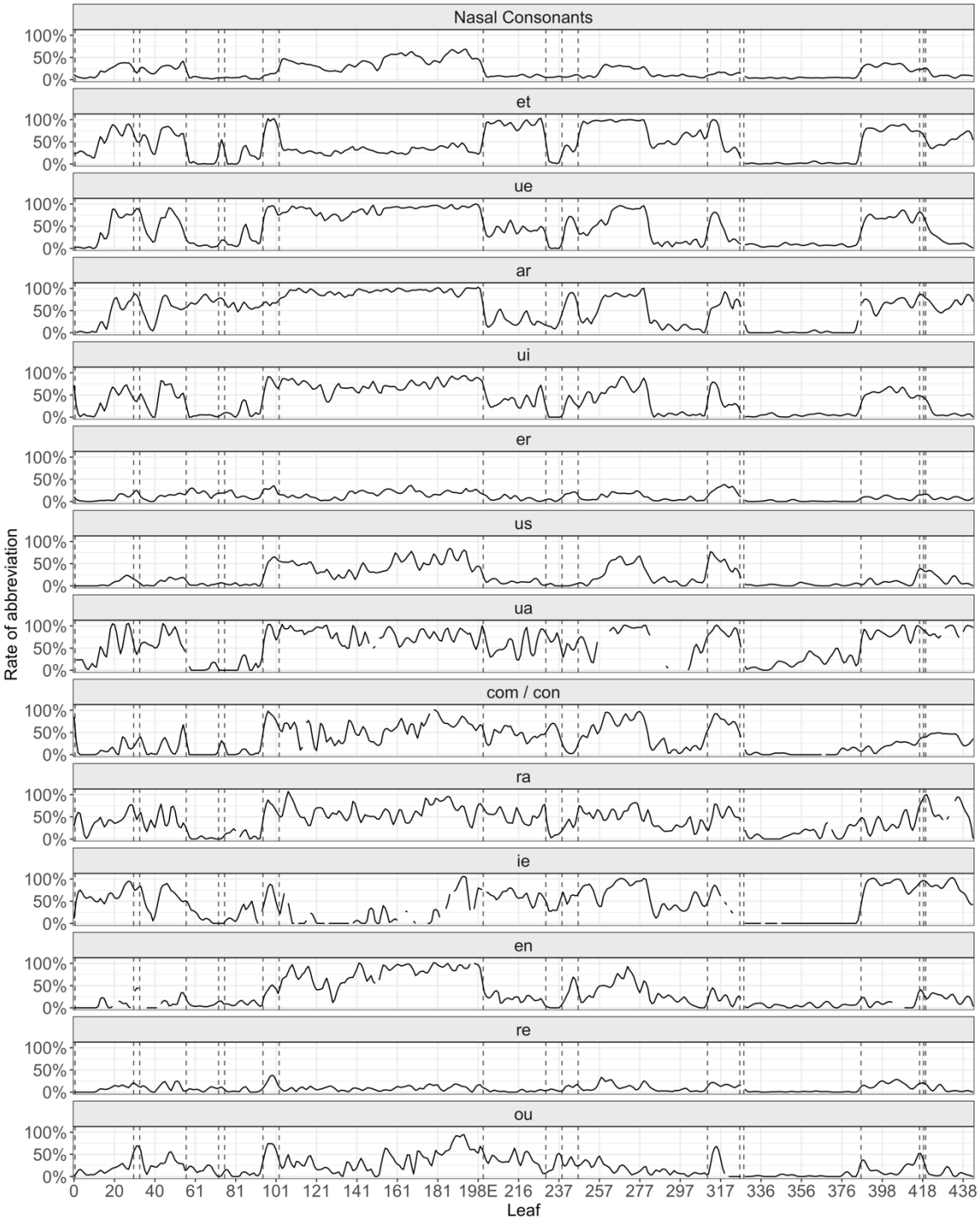
On one leaf, 100% of the possible candidates may have been abbreviated, while on another, no such abbreviation may have occurred. There are moments when the rates of abbreviation appear to fluctuate, and others

where they are relatively static. There are also moments where the abbreviatory practices appear to change sharply, with the rates of abbreviation for a given feature shooting up or down. These, it appears, often align with moments of scribal juncture, where one scribe, with one unique abbreviatory profile, hands over to another, who in turn has a different profile. By overlaying the plot with the scribal junctures we identified through traditional palaeographic techniques in Chapter 1, as presented in Figure 2.39, the results are stark.

The vast majority of occasions where a clear moment of change occurs in the use of abbreviation line up closely with a change in scribe, such that we may say that scribal preference plays a significant role in the abbreviatory choices in fr24432. It also appears possible to identify general trends in a given scribe's abbreviatory usage, such as the fact that hand 6 (responsible for ff. 103ra–198G^{vb}) appears to make greater use of most forms of abbreviation compared to their peers, while abbreviation is relatively scarce in the work of hand 10 (ff. 328(bis)^{ra}–385^{rb}).

[24] To smooth the data to the point where they can be interpreted clearly, I apply Locally Estimated Scatterplot Smoothing (LOESS) to the data. While this improves the readability of the data, it also means that in the most extreme shifts between rates, these can appear to change over a number of folios, when in fact the shift is from one folio to the next. Gaps in the data appear where there are no instances of a given abbreviated sequence on that particular folio.

Figure 2.39 – rates of abbreviation for the fourteen most frequent abbreviated sequences per leaf in fr24432, with scribal junctures.



Rather than basing conclusions on the subjective interpretation of a plot, however, we will achieve more reliable results through statistical methods, examining whether the scribal junctures do indeed line up with moments of significant change in the rates of abbreviation. This will then allow us to reach concrete conclusions about the work of individual scribes.

In order to identify moments at which the uses of abbreviation change significantly, I first calculate a standard deviation for the whole manuscript per abbreviated sequence. Then, two averages are calculated per sequence and per leaf, one looking backwards five leaves and one looking forwards five leaves (both excluding the leaf in question), in order to determine a reliable value for the abbreviatory trends present before and after a given leaf, without straying too far from it. The rate of abbreviation on each leaf is then compared to these averages, through a calculation of how many standard deviations away from each mean the observed rate is. The difference from the backward mean minus the difference from the forward mean is then calculated as an indicator of which mean the observed rate is closest to, with positive values indicating a greater affinity with the forward mean than the backward mean, that is to say that the leaf in question displays a greater difference from the leaves that come before it than the leaves that come after it. These values are calculated for all grapheme sequences of interest and combined into a single figure per leaf. Differences in the rates of abbreviation for a given leaf are said to be significant if this total figure is greater than or equal to half the number of the grapheme sequences under examination. That is, the abbreviatory practices of a given leaf can be said to indicate a shift in practice if, on average, each grapheme sequence of interest has a difference from the backwards-facing mean which is at least 0.5 standard deviations greater than the difference from the forwards-facing mean. In practice, this number is unlikely to be consistent across all the abbreviated sequences being considered, since scribes may vary in the extent to which their practices differ from others in a non-linear way, depending on the specific abbreviatory process in question.[25]

This algorithm returns 13 leaves for which a change in practice is identified, namely ff. 1, 42, 57, 95, 103, 199, 231, 247, 258, 282, 311, 328(bis), and 388. Comparing these to the scribal

[25] Within the leaves for which a change is identified, the algorithm often outputs a number of runs of consecutive or near-consecutive leaves, where a significant change in abbreviatory usage causes a shift to be identified even after the initial leaf at which the change occurred. I therefore exclude all but the first leaf of such sequences.

junctures which we identified in fr24432 using traditional palaeographic methods, we note that 10 align exactly with a shift in scribe, reinforcing the sense that the most significant factor determining variation in abbreviatory uses across the pages of fr24432 is the preferences of the scribes involved in its production. It is therefore possible to predict the location at which a scribal juncture may occur based on a statistical examination of the rates of abbreviation observed within the manuscript.

This statistical approach on its own, however, is insufficient to identify every moment of scribal juncture, since there are 9 shifts in scribe which go undetected by the algorithm. Perhaps unsurprisingly, given it considers 5 leaves either side of the focus leaf, the algorithm is ineffective at determining the position of scribal junctures where an individual scribal stint lasts only a few pages. Thus the brief intrusion of hand 3 within a broader section by hand 2, which appears on ff. 29^{vb}–32^{vb}, is missed at both of the stint's junctures, as is the intervention of hand 4 within a section by hand 3 on ff. 73ra–76ra. Further short stints by hand 3 (ff. 239ra–246^{vb}), hand 9 (ff. 327ra–328^{vb}), and a longer section by hand 11 (ff. 416va–443^{vb}), itself interrupted by hand 2 (ff. 419ra–^{vb}), are also missed. In all cases except the final stint by hand 11 (ff. 420ra–443^{vb}), this can be explained by the fact that the stint in question is shorter than the 11 leaves used by the algorithm to reach its conclusions. In the case of the final stint, the lack of identification of the scribal junctures can perhaps be explained by the fact that while scribes 2 and 11 do appear to differ in their abbreviation of a number of grapheme sequences, this difference is not quite as significant as that which is seen at other junctures. Thus, while it is not always able to identify scribal junctures which occur within a manuscript, this statistical approach nevertheless offers a useful tool, particularly as a means of confirming, through concrete, reproducible means, the conclusions of traditional, palaeographic approaches which are often somewhat more subjective.

Notably, there are three moments identified by the algorithm which do not line up with scribal junctures as identified through our earlier analysis. The algorithm identifies f. 42 as sharing a much greater affinity in its rates of abbreviation with the pages that follow it than with the pages that precede it, and this is also visible in Figure 2.39, in which we observe a sharp increase in the rates of abbreviation of ⟨et⟩, ⟨ue⟩, ⟨ar⟩, ⟨ui⟩, ⟨ua⟩, ⟨com⟩ and ⟨con⟩, and ⟨ie⟩ in particular. The hand used across this shift is largely consistent, such that it does not appear to be the case that this leaf marks a scribal juncture. What is particularly notable, however, is the fact that early on f. 42ra we shift from the end of one text (item 11, *Le dit des trois chanoinnes*) to the beginning of another (item 12, *Le lai de l'oiselet*). Both texts are written in verse, with the former consisting of alexandrine lines split on the page into six-syllable hemistichs, and the latter in octosyllables. There are therefore two feasible explanations for this shift, either that the scribe changed their abbreviatory practices as a result of the shift in verse form, or because they were influenced by an exemplar copy of *Le dit des trois chanoinnes* in which abbreviation was less frequent, before turning to another exemplar (or the work of a different scribe within the same exemplar) in which abbreviation was more common.

A second shift which is identified by the algorithm but which is not the result of the presence of a scribal juncture is that found on f. 258. Again, the shift in usage is visible in Figure 2.39, through an increase in the abbreviation of the nasal consonants, as well as the sequences ⟨ar⟩, ⟨ui⟩, ⟨us⟩, ⟨com⟩ and ⟨con⟩, ⟨en⟩, and ⟨re⟩ in particular, and a reduction in the use of ⟨ie⟩. Again, there is no evidence of a change of scribe here, but there is a change in text, with the alexandrine *Le dit du buef* (item 61) giving way at the end of f. 257va to the octosyllabic item 62, *Le dit des rues de Paris*. The comparable shift in metre to that observed above for f. 42 reinforces the possibility that poetic form may play a role in the rates at which the scribes of fr24432 chose to abbreviate certain grapheme sequences, a question which we will examine in further detail below. It is also significant that in both cases examined so far, the shift has been from a

text which, as we will argue in Chapter 3, can be attributed to the same author. This reinforces the possibility that the influence of exemplars may be at play here, given the reasonably comparable rates observed for *Le dit des trois chanoinnes* and *Le dit du buef*, suggesting that these two texts by the same author may have been taken from a single exemplar manuscript with relatively constant rates of abbreviation, and were placed in disparate positions within the base manuscript of fr24432, alongside texts from exemplars with different rates of abbreviation.

The possible impact of exemplar manuscripts must be tempered, however, by the fact that there are no clear shifts in the way that these texts themselves are presented. Despite the presence of shifts across fr24432 in the kinds of *incipits* and *explicits* used for each text, which do not line up with changes in scribe but rather from apparent shifts in exemplar, both items 11 and 12, as well as items 61 and 62, have comparable *incipits* and *explicits*. Both pairs are also comparable in their use of *litterae notabiliores*, and as such we must at the very least conclude that the influence of exemplars is not the only possible explanation for the shift in abbreviatory practice at both ff. 42 and 258.

The final shift in abbreviatory usage identified by the algorithm is perhaps the most significant. At f. 282, as can again be identified from Figure 2.39, there is a marked decrease in the rates of abbreviation of almost every grapheme sequence which we have considered. Despite this significant shift in abbreviatory usage, which resembles the shift between hands 6 and 7 on f. 199, there is once again little palaeographic evidence for a shift in scribe at f. 282, rather the explanation for the shift in usage appears once again to come from one or other of the form of the texts being copied or their source, with the former appearing more likely in this case. In the final column of f. 282^{v}, we find the end of *Le dit du vilain despensier* (item 70), a very short poem which began only on the recto side of the same leaf, which is the last in a sequence of three similar poems, beginning on f. 279va. The text which begins on f. 282^{vb}, *Li*

livres des songes Daniel et les songes Macrobe (item 71), meanwhile, is markedly distinct in form, being a prose text which runs for some 21 leaves. It is the distinct forms of these two texts, therefore, one being in verse and the other in prose, which appears to offer the best explanation for the significant shift in abbreviatory usage.

From this evidence, it is clear that there are a number of different factors which can influence the rates at which individual grapheme sequences are abbreviated within fr24432, most notably the copying scribe, but also the form and metre of the text being copied. Let us therefore consider these factors individually across the whole manuscript, to identify the extent to which they can be seen to influence abbreviatory practice.

2.4.3 Factors affecting abbreviatory usage

Scribe

As we identified within Figure 2.39, the preferences of the scribe who copied a given text appear to have a significant impact on the rates of abbreviation that occur within it. It ought therefore to be possible, for the scribes of fr24432, to add to our sense of each scribe's abbreviatory profile by determining the rates at which they abbreviate the most common sequences when unaffected by external factors. In order to do so, we must look only at texts with a single scribe, since there appears to be a significant chance that when two scribes collaborate on a single text, one may be influenced by the abbreviatory practices of the other. The scribes we consider here who work on such single-scribe texts are therefore hands 2, 3, 5, 6, 7, 10, and 11.[26]

[26] Excluding hands 1, 4, and 9 in particular from the analysis appears to be logical anyway, given they copy only a few columns each.

Figure 2.40 – rates of abbreviation for the fourteen most frequent abbreviated sequences in the work of select scribes within fr24432.

Figure 2.40 displays rates of abbreviation for the most commonly abbreviated sequences in the work of these scribes, and allows us to directly compare their preferences. On the whole, hands 2, 5, 6, 7, and 11 tend to abbreviate a great deal, while hand 10 in particular

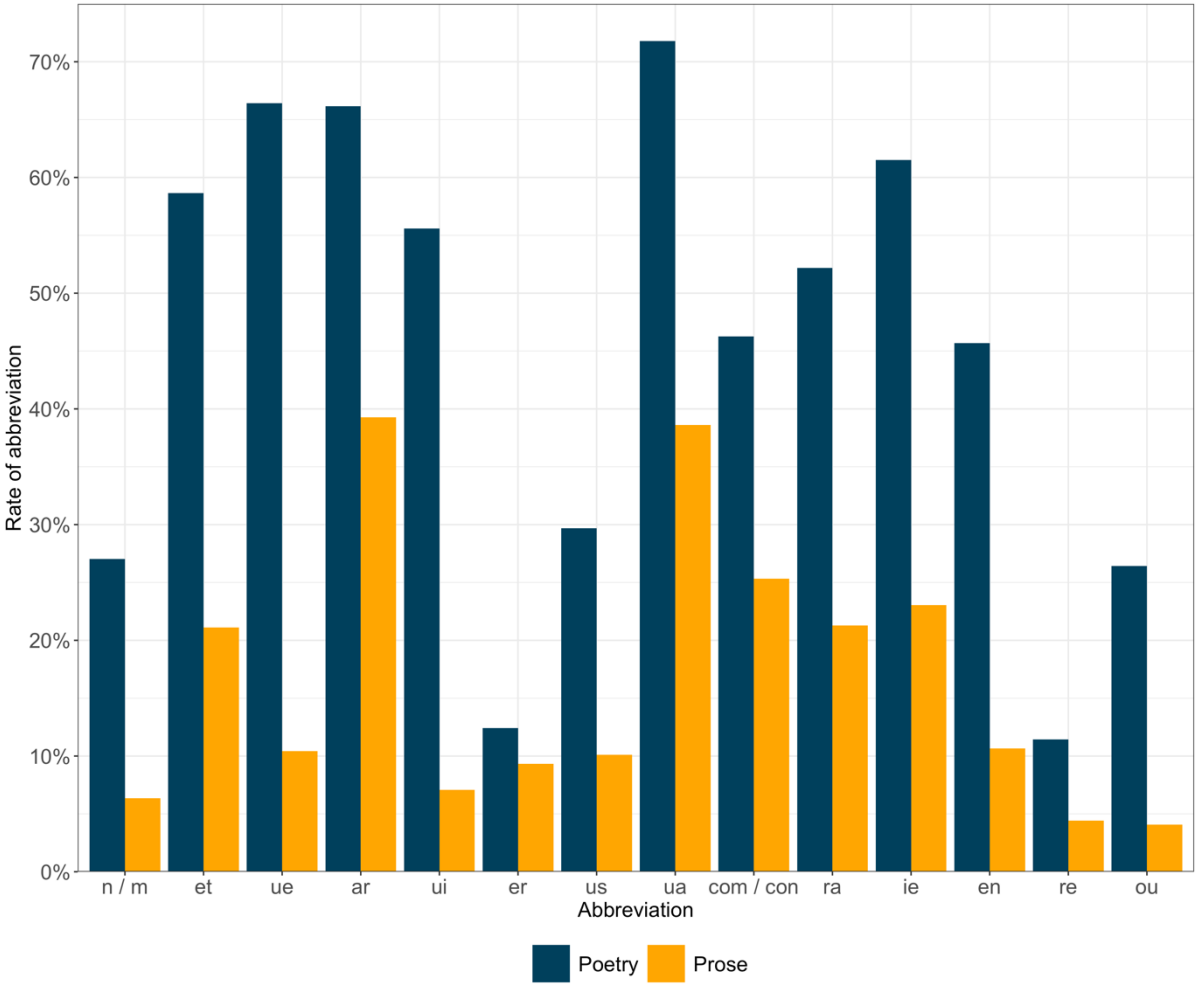
abbreviates very little, but the data paint a more complex picture than simply a general preference for or against abbreviation. Hand 11, for instance, displays low rates of abbreviation for nasal consonants, ⟨ue⟩, ⟨ui⟩, ⟨er⟩, ⟨us⟩, ⟨en⟩, ⟨re⟩, and ⟨ou⟩, but much higher rates for ⟨et⟩, ⟨ar⟩, ⟨ua⟩, ⟨ra⟩, and ⟨ie⟩ in particular, while the rates of abbreviation displayed by hand 5 are amongst the highest of any scribe for all sequences except the nasal consonants. The scribe copying a particular text therefore has a clear impact not simply on how often abbreviation occurs, but rather which forms are abbreviated more than others.

Form

The kind of text being copied, as we have seen, also plays a significant role in the rates at which abbreviation occurs. In her examination of abbreviation in a number of medieval French manuscripts, Geneviève Hasenohr suggests that abbreviation appears to be more common 'dans les copies d'œuvres proprement littéraires, épopées d'abord, romans ensuite, que dans les manuscrits contenant des pièces satiriques ou des traités didactiques' (2002: 83). Within fr24432, of course, we find texts matching all of these descriptions, often more than one at once, and so an assumption that manuscripts and their texts can be grouped into such straightforward categories is far from unproblematic. It is however possible to make a clearer distinction by considering a given text's form, rather than its content. That is, we can compare the rates of abbreviation found in poetry to those found in prose, in order to identify and distinctions which appear between them. Figure 2.41 displays the rates of abbreviation for the most frequently abbreviated sequences in fr24432, when divided according to whether the text in which it is found is written in poetry and prose.[27]

[27] I have therefore excluded abbreviations found in paratextual material, as well as Rutebeuf's *L'erberie* (item 9), which is in mixed verse and prose form.

Figure 2.41 – rates of abbreviation for the fourteen most frequent abbreviated sequences in single-form texts of fr24432, for poetry and prose.



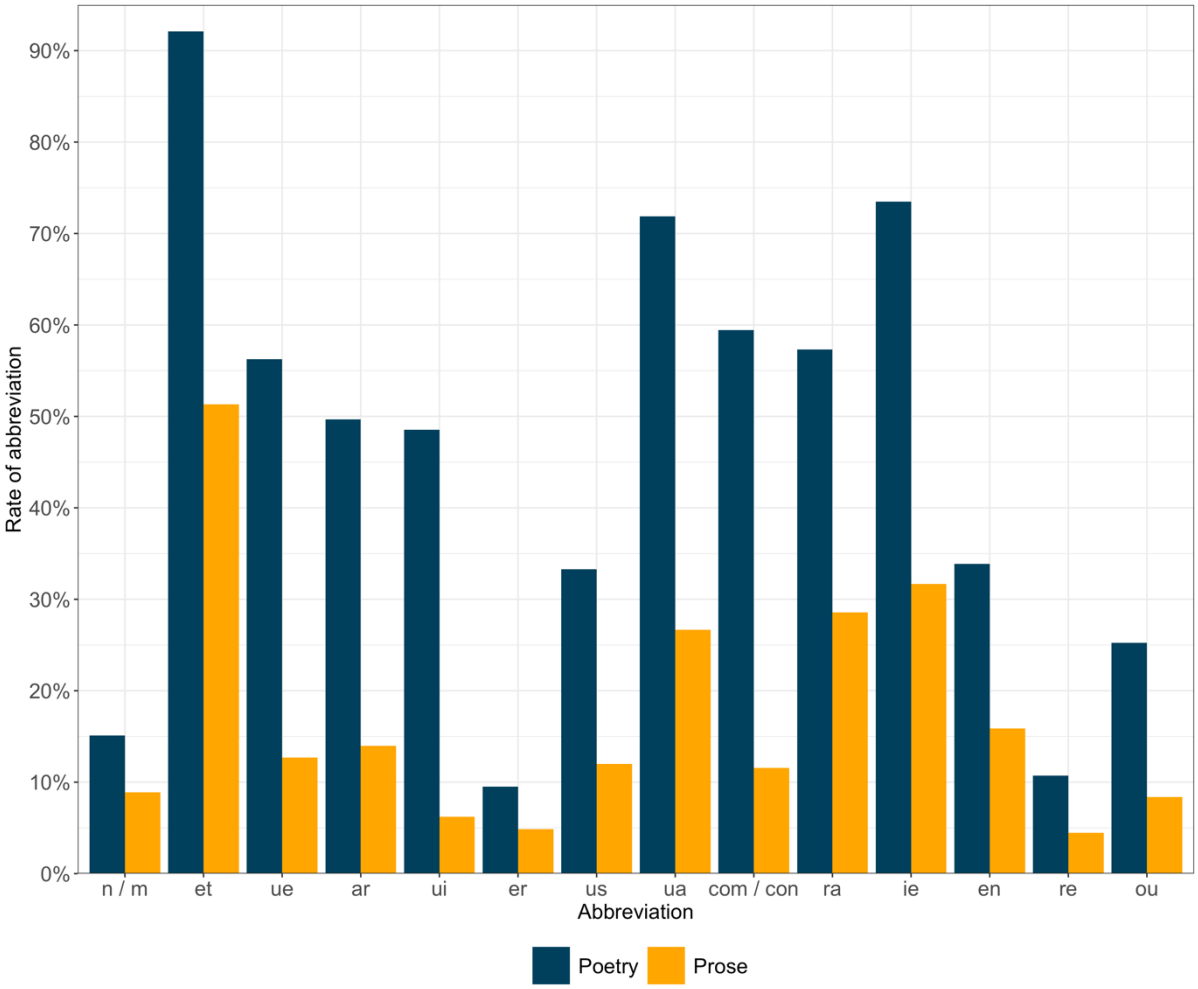
From the data, we observe that within fr24432, poetry contains consistently higher rates of abbreviation than prose, and these differences are statistically significant, with a p-value test giving a result of $p = 0$ for all pairs. While in some cases the rates of abbreviation do not differ particularly dramatically (as in ⟨er⟩), in some cases, as in ⟨ue⟩, the rates observed in poetic texts are several times greater than those observed for prose.

Given that, as we saw on f. 282^{v}, a shift between poetry and prose can trigger a shift in abbreviatory usage irrespective of scribal preferences, it appears likely that these differential rates are indeed conditioned by form, rather than scribal preference. By looking at the work

of hand 7, the only scribe in fr24432 to copy entire texts of substantial length in both verse and prose, we are able to confirm this statistically.[28]

Hand 7 is the sole copyist of some 15 texts,[29] of which all but one, *Li livres des songes Daniel et les songes Macrobe* (item 71), are verse texts, in a mixture of hexasyllable, octosyllable, and alexandrine. *Songes* itself is of substantial length, containing some 85 columns of text, and therefore provides ample candidates for abbreviation, allowing us to make direct comparisons with the poetic texts copied by hand 7.

Figure 2.42 – rates of abbreviation for the fourteen most frequent abbreviated sequences in the work of hand 7, for poetry and prose.



[28] Hands 3 and 5 also copy two prose texts, items 25 and 27, but both are too short to provide a useful sample.
[29] Items 53–56, and 61–72.

The data for this single-scribe example once again display much higher rates of abbreviation for poetry than for prose. With the exception of ⟨ua⟩, for which an insufficient number of candidates exists in *Songes* to provide a large enough sample ($p = 0.07107$), these distinctions in rates are statistically significant. The example of hand 7 therefore allows us to extrapolate and state with some certainty that, in fr24432 at least, individual scribes use abbreviation at much higher rates in poetry than they do in prose.

This goes against the established assumptions surrounding scribes' motivations for abbreviation, particularly concerning the preservation of writing materials, which would suggest that prose should contain higher rates of abbreviation than poetry. Since the lengths of written lines of prose are not fixed, it is possible to save great quantities of parchment and ink, particularly for longer texts, by sustaining high rates of abbreviation, and yet we find from these data that such material considerations did not have an impact on abbreviatory usage.

There is, however, an alternative material explanation of significance. Precisely because poetry has fixed line lengths, we may in fact hypothesise that such texts necessitate greater rates of abbreviation than prose. When writing poetry throughout most of fr24432, the aesthetic imperative for the scribes appears to have been overwhelmingly to ensure that each line or half line of poetry could fit within a single line of text.[30] As such, scribes may well have felt that it was better to be safe than sorry, making greater use of abbreviation so as to make sure that this imperative would never be broken. This could go some way to explain the examples mentioned by Hasenohr (2002: 91), where abbreviation appears in a line of verse that has nevertheless not reached the edge of the column, and which the scribe has then filled in by extending the final letters. The scribes in question, when deciding whether to abbreviate

[30] With the exception of item 72, a collection of *Serventois*, which are presented as concatenated text with a *punctus* between each line of verse.

certain sequences, were balancing the requirements of a first-order aesthetic imperative—that verse lines must not overhang their allotted space—with a second-order one—that columns should be fully justified—and occasionally went too far in adherence to the former.

Metre

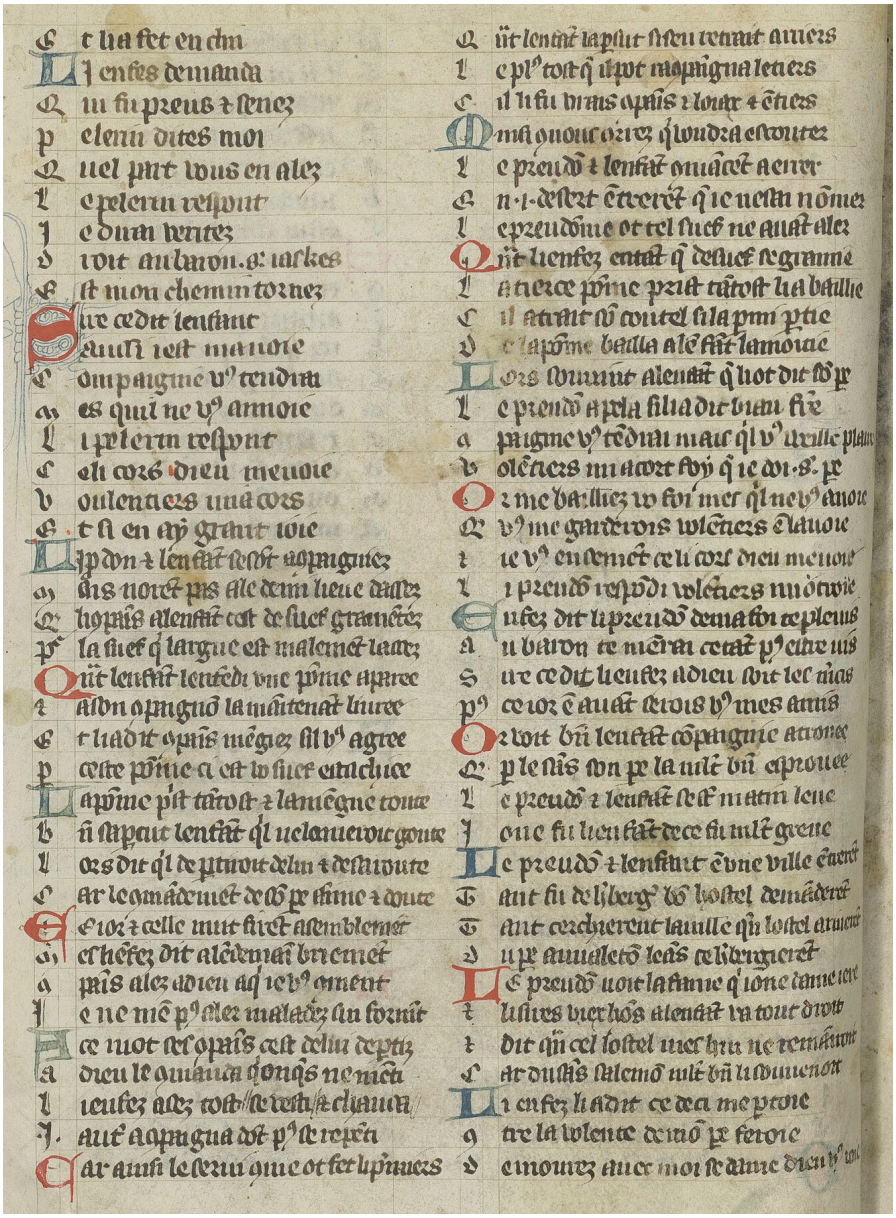
If this material explanation were to be true, then by extension we could expect to see greater rates of abbreviation in lines of poetry containing a greater number of syllables than those where the syllable count is lower, displaying scribal awareness of the more complex aesthetic difficulties they faced. A clear example of this can be found within fr24432 on f. 54^{v}. This folio, copied by hand 2, contains part of a text in alexandrine verse, which for the most part is separated into lines of 6 syllables each, as is common for alexandrine poetry throughout fr24432. For reasons which remain unclear, perhaps with the aim of material preservation, halfway down the first column of f. 54^{v}, the scribe decides to change this approach, and includes full lines of 12 syllables in each line of text.[31] As a result, they have to deploy a number of strategies in order to reduce the width of each line, including a general narrowing of the script, visible from the facsimile image in Figure 2.43, as well as an increase in the use of abbreviation.

Taking a simple quantitative approach, we find that in the 21 lines of f. 54^{v} which are written as full alexandrines, there are 81 instances of abbreviation, appearing at a rate of roughly 4 per line. In the preceding 21 lines, however, in which alexandrines are split into hemistichs, we find just 6 instances of abbreviation, that is, abbreviation appears on average only once per 3.5 lines. The clear distinction between these sections, despite being the contiguous work of a single scribe copying a single text, thus gives a strong indication that hand 2 saw

[31] By the top of f. 55^{r}, the scribe has clearly decided not to continue with this approach, and so gradually returns to dividing alexandrine lines into two separate lines of text, at first only when a line has a particularly high character count, and later more consistently.

abbreviation as one of the possible strategies for condensing lines of poetry in order to meet the demands of the aesthetic and material constraints under which they worked.

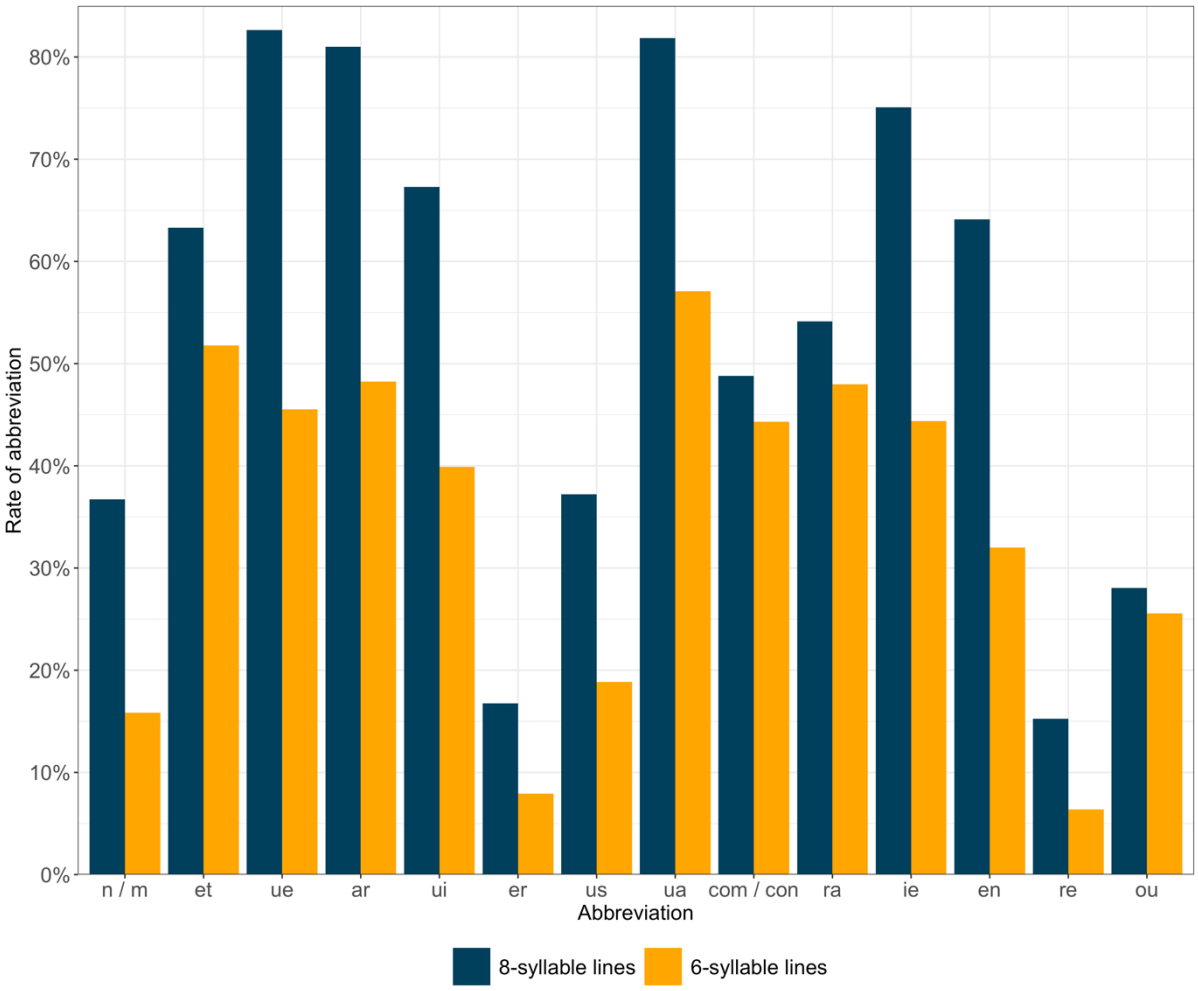
Figure 2.43 – f. 54ᵃ of fr24432.



To take a more holistic approach across the whole manuscript, let us compare the rates of abbreviation used in lines of 6 syllables (including split alexandrines) compared to those of 8. If we do so then we find, as presented in Figure 2.44, that for each of the abbreviated sequences we have considered, there is more abbreviation to be found in 8-syllable lines than

6-syllable ones,[32] thereby appearing to confirm our hypothesis that, for fr24432 at least, abbreviation is used in lines of poetry as a way of ensuring that the line will be narrower than the width of the column.[33]

Figure 2.44 – rates of abbreviation for the fourteen most frequent abbreviated sequences in texts in fr24432 presented in 8- and 6-syllable lines.



Position in line

If condensing a line were a scribe's sole motivation for an increased use of abbreviation, then we might expect that abbreviation would become increasingly common as a line progressed,

[32] The distinction in rates is statistically significant for all sequences except ⟨ou⟩ ($p = 0.17394$).
[33] This variation between 8- and 6-syllable lines is also evidence against the suggestion that abbreviation may be more frequent in poetry than prose because abbreviated rhyming syllables can be more easily deciphered using surrounding context. We may also wish to consider whether there is a difference in the rates of abbreviation displayed in strophic texts, as opposed to non-strophic ones, but there are unfortunately insufficient examples of strophic and non-strophic texts of the same metre in fr24432 to enable us to answer this question.

with a scribe using more and more abbreviation if it became apparent that the line in question would otherwise not fit within the column. This assumption, that abbreviation is more common at the ends of lines, is well established within scholarly approaches to manuscript abbreviation, but the evidence for such a claim appears to be almost entirely anecdotal.[34]

Through an examination of the written lines of fr24432, we can identify the position at which each abbreviation appears within its containing line. These data are represented in Figure 2.45, with each abbreviation assigned to a class based on its percentage position in a given line, an approach that allows us to compare the use of abbreviation across the vastly different line lengths within the manuscript.[35]

Figure 2.45 – number of abbreviations at a given position in a line across fr24432.

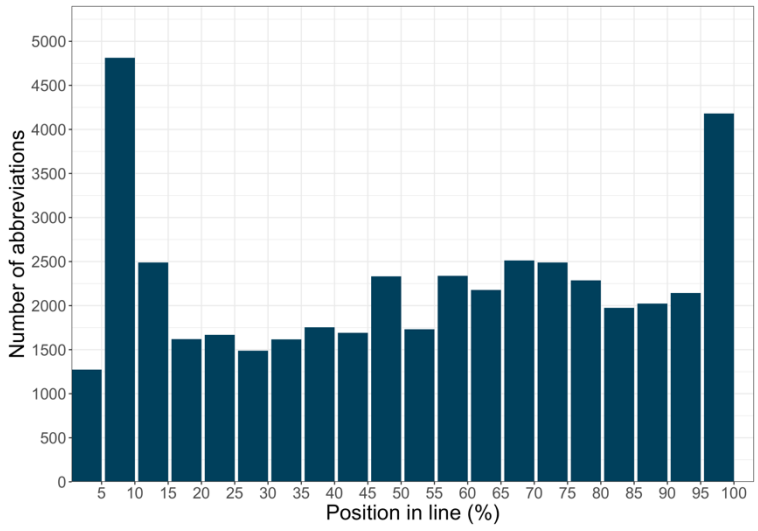


Figure 2.45 does arguably go some way to support the widespread assumption that abbreviation is common at the ends of lines, given that a large number of abbreviations are found in the last 5% of the line, in practice usually the last or second to last character, which

[34] See, for example, Daly's (1950: 213) use of numerical counts of abbreviations at the end of the line as a means of scribal identification in a ninth-century Latin manuscript.
[35] I count spaces between words as distinct characters, given that they take up width within a line just like any other character.

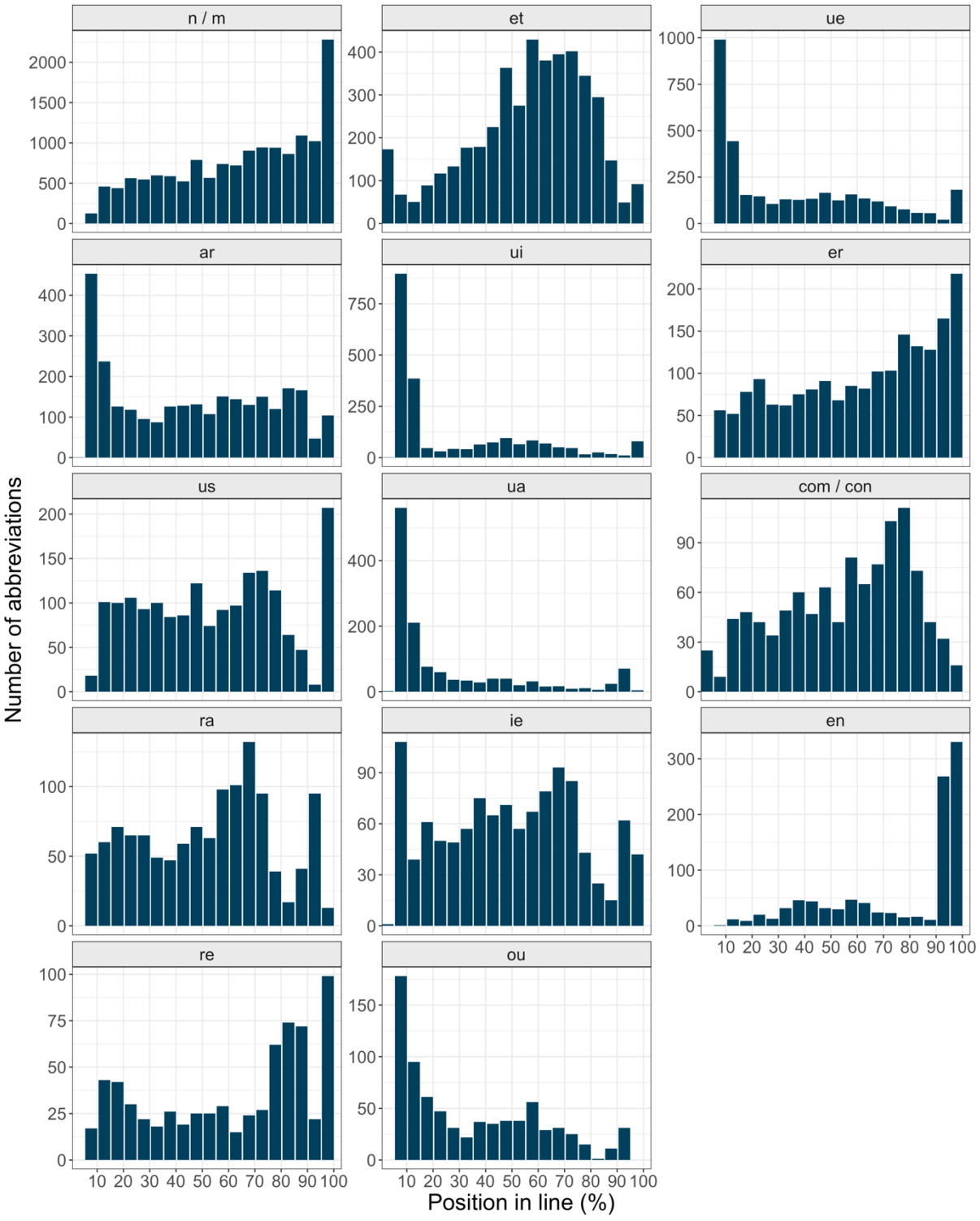
is consistent with scribes choosing to abbreviate sequences once they realise that an unabbreviated form would result in an over-length line.

What is most striking about these data, however, is what they tell us about the other end of a line, namely that the most common position for abbreviation to occur is between 5 and 10% of the way through the line, which in most cases is after the first character. Scribes would certainly have been able to determine from their exemplars whether the line they were copying would be more difficult to fit into their column, and it could therefore be the case that the greater prevalence of abbreviation at the beginning of a line is the result of scribes pre-empting a more challenging line by abbreviating it at the earliest opportunity. Evidence to support this claim comes from the fact that if an abbreviation occurs at the beginning of a line as the second character, then there is a higher rate of abbreviation per character across the rest of the line (3.47%) compared to across the rest of the manuscript (2.51%, $p = 0$), suggesting an awareness on the part of the scribe that abbreviation will be necessary in the line that they have begun. Notably, there is also a comparable trend for lines containing an abbreviated sequence in final position, with the rate of 2.85% elevated compared to the average in the rest of the manuscript (2.12%, $p = 0$), suggesting that even at those moments when abbreviation proves necessary at the end of a line, scribes have made greater use of abbreviation elsewhere in the line in an attempt to condense it before the end.

As with the overall rates of abbreviation, it is worthwhile considering the positions at which individual abbreviated sequences appear within a line, in order to determine where certain sequences are most frequent. Looking at the data in Figure 2.46, we observe that some 6 sequences, namely ⟨ue⟩, ⟨ar⟩, ⟨ui⟩, ⟨ua⟩, ⟨ie⟩, and ⟨ou⟩ appear with greatest frequency at the beginning of a line, while a further 5, specifically the nasal consonants, ⟨er⟩, ⟨us⟩, ⟨en⟩, and

⟨re⟩ are most frequent at the end of a line. The sequences ⟨et⟩, ⟨com⟩ and ⟨con⟩, and ⟨ra⟩, meanwhile, are most frequent towards the middle of a line.

Figure 2.46 – number of abbreviations at a given position in a line for the fourteen most frequent abbreviated sequences within fr24432.



Many of these tendences can be explained through syntactic means, such as the fact that ⟨ue⟩, ⟨ui⟩, and ⟨ua⟩ most commonly appear after ⟨q⟩, while ⟨ar⟩ is most common after ⟨p⟩, and ⟨que⟩, ⟨qui⟩, ⟨quant⟩, and ⟨par⟩ are common clause-initial particles with which it is possible to begin a line, particularly a line of poetry. It is therefore likely that scribes abbreviated these sequences frequently within such words, but that these are simply more common at the beginning of a line, rather than choosing to abbreviate them for any material reasons.

⟨en⟩ is also particularly striking, in that the vast majority of its appearances are within the last decile, most likely as a result of the fact that, as we have seen, this sequence is most commonly abbreviated as part of the sequence ⟨ment⟩ and its variants, which is a common structure at the rhyme.[36] In this case too, then, the tendency for abbreviations of this kind to appear towards the end of a line is predominantly due to the position that ⟨en⟩ tends to occupy in a line of verse, although in this case it is certainly possible that material factors are also at play.

A number of these sequences also show statistically significant correlation with their position in the line, such that we can say that they become more or less frequent as a line goes on. The nasal consonants ($r = 0.82110$, $p = 0.00001$), as well as ⟨er⟩ ($r = 0.86364$, $p = 0.00000$), ⟨en⟩ ($r = 0.55668$, $p = 0.01079$), and ⟨re⟩ ($r = 0.59240$, $p = 0.00592$) are much more likely to occur towards the end of a line than towards the beginning. Meanwhile, ⟨ua⟩ ($r = -0.46202$, $p = 0.04028$) and ⟨ou⟩ ($r = -0.54449$, $p = 0.01306$), on the basis of their negative correlation coefficient, diminish in frequency throughout a line.

Thus, while little can be said about the impact of the position in a line on the frequency of abbreviation other than at the very beginning or end, when we consider individual abbreviated sequences, the significance of where such a sequence appears becomes evident.

[36] For more discussion on the frequency of certain structures at the rhyme, see the p. 189 and following.

Line length

A further possible contributing factor determining the frequency with which abbreviation occurs is the length of the line in which it appears. Interestingly, however, it is not straightforwardly the case that longer lines lead to a greater rate of abbreviation. There is a very weak but statistically significant positive correlation between a line's length and the number of abbreviated sequences it contains ($r = 0.03206$, $p = 0.00000$), such that longer lines do on the whole appear to coincide with a greater number of abbreviations, but line length alone does not appear to offer a sufficient explanation. Indeed, comparing the top decile of lines across fr24432 (at least 30 characters in length) to the bottom decile (at most 17), we find a remarkable difference in rates, with a rate of occurrence of abbreviation per character of 1.65% for the top decile and 3.59% for the bottom decile ($p = 0$). That is to say that in the shortest lines within fr24432, the rate at which abbreviation appears per character is in fact more than twice the rate observed in the longest lines. While material factors therefore appear to have played some role in decisions to include abbreviation, in many cases it appears likely that the inclusion of a long line or a short line was as much an aesthetic and a textual choice as a material one.

Position on page

A similarly tempting explanation for the use of abbreviation in a given position could be that as scribes came closer to the end of a page, perhaps with a certain point in mind at which they wanted to finish the page, they would use more abbreviation in order to condense their text and ensure that they achieved this goal. In order to examine this suggestion, we must of course focus on pages containing only prose texts, since as we have seen, there is a significant aesthetic imperative across fr24432 to keep lines of poetry in the same written line, regardless of whether this results in an inefficient use of parchment.

Figure 2.47 – number of abbreviations at a given position on a page across the prose texts of fr24432.

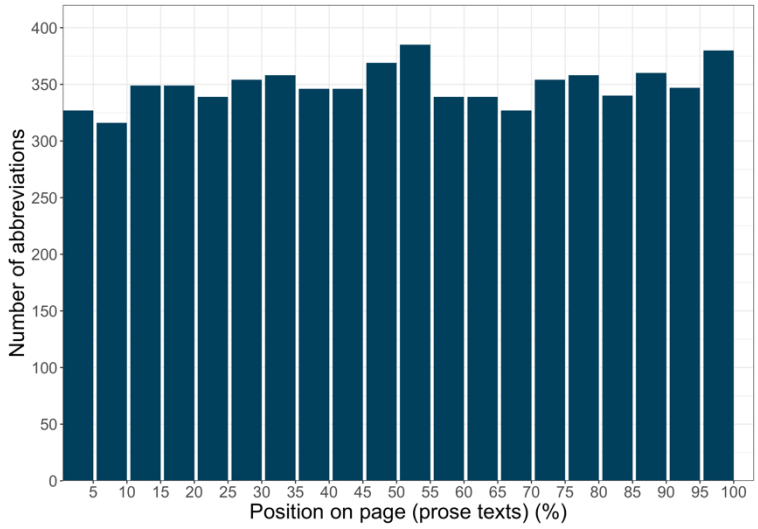


Figure 2.47 presents the ventile position on a page of those abbreviated sequences within fr24432 which appear on pages containing only prose. On the face of things, the data appear to be relatively static, but there is a positive correlation between the number of abbreviations and their position on a page (r = 0.40290), which would suggest that abbreviation becomes more frequent towards the end of a page. In this case, however, the p-value is insufficient to determine statistical significance (p = 0.07818), meaning that we are unable to confirm this link with any certainty. Notably, however, and despite our initial exclusion of poetic texts, a similar correlation (r = 0.53678) appears across the manuscript as a whole, and this time the correlation is statistically significant (p = 0.01468), such that even though there is a limited chance in fr24432 of increased abbreviation in poetry reducing the number of lines required by a given section of text, the scribes do nevertheless appear to have made greater use of abbreviation as they approach the end of a page.

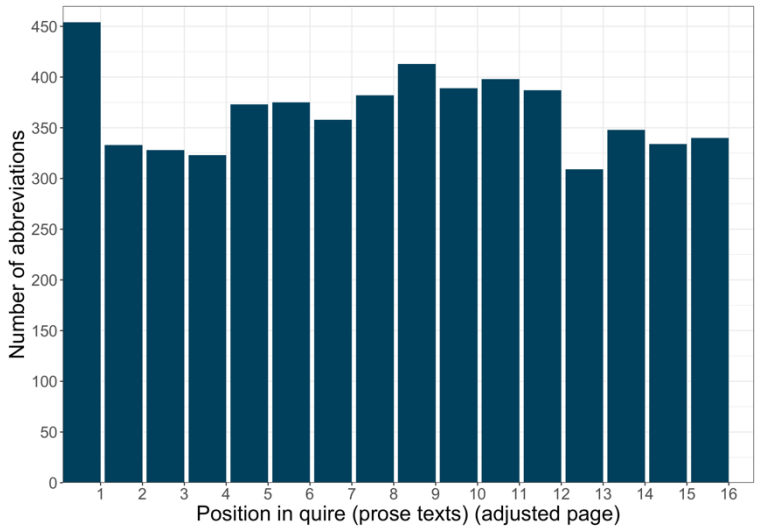
Position in quire

A final position of possible significance is that of an abbreviated sequence within a quire. Just as with a page, one might suspect that scribes would make greater use of abbreviation towards the end of a quire, if they had to complete a certain quantity of text by the end of

their quire. Again, it appears most logical only to examine those quires which contain exclusively prose texts, since their layout is most malleable.

Figure 2.48 presents the number of abbreviations at a given position in a quire across the prose texts of fr24432. Here the data are divided into 16 classes, on the basis of the fact that most quires within the codex consist of 8 leaves, and thus 16 sides in total. For quires where this is not the case, I have calculated an adjusted position for each abbreviated sequence, based on its relative proximity to the beginning and end of the quire in which it appears.

Figure 2.48 – number of abbreviations at a given position in a quire across the prose texts of fr24432.



The data here allow us to say with some certainty that in fr24432 it is not the case that abbreviation increases as a scribe moves towards the end of their quire, since the calculated correlation value is in fact negative (-0.23122), which would suggest that the rate of abbreviation decreases as a scribe approaches the end of a quire, and the p-value for correlation (0.3889) is well above the threshold for statistical significance anyway.[37] Of note, however, is the somewhat elevated number of abbreviations that appear at the very beginning of a quire. If we take the start of a quire to be the start of a period of work, then it is

[37] The data across the manuscript as a whole tell a similar story, giving a weak negative correlation which is not statistically significant ($r = -0.19315$, $p = 0.4735$).

interesting that abbreviation should be most common on the first page of a quire. It could suggest that elevated levels of abbreviation were associated with good aesthetic or scribal practice, and that scribes made an effort to abide by this practice when beginning their work, but that—just as is often observed with the precision of written forms and the avoidance of errors—scribes reverted to a more comfortable standard of practice after the beginning of their work.

2.5 Conclusion

This chapter set out to display the utility of the analysis of manuscript abbreviation as part of the manuscript scholar's toolbox, and our case study has shown that such an approach can reveal facts about a manuscript's compilation that both strengthen existing codicological and palaeographic methods and bring about new conclusions. We have seen that abbreviatory features can be used to identify moments of change within a manuscript, which most often appear to correlate with changes in scribe, but which can also be indicators of other shifts, such as a change in form or metre between two texts, or even the exemplar from which a given text is copied.

There is an incontestable value to taking scribal abbreviation as an object of analysis, rather than simply a challenge to be overcome when reading a manuscript, and our functional approach allows us to approach manuscript abbreviation with much more clarity than has previously been possible. Much remains to be done, however. On its own, an analysis of abbreviation can reveal facts about manuscripts which are otherwise difficult to identify, but it is by combining this approach with other methods that its utility really comes to the fore. Future work on combining analysis of abbreviation with better-studied scribal features such as orthography is therefore necessary to explore how this statistical approach can be strengthened further.

In short, the maligned status of abbreviation within manuscript studies, particularly the avoidance of the question of what exactly is abbreviated rather than how, has led to a dearth of understanding of the processes that bring about abbreviation, and how it varies from scribe to scribe and from manuscript to manuscript. It is therefore necessary to redress this balance, through methods only recently made possible through the advent of the digital humanities. By gaining a clearer understanding of the role that abbreviation played in the medieval writing system, we can learn a great deal about the texts and objects which contain it.

3. A la recherche d'un auteur perdu

Digital, statistical, and codicological approaches to a case of uncertain authorship in fr24432.

3.1 Introduction

Medieval authorship is, to say the least, a fragile concept. There is so often little to no surviving evidence about who wrote a given text, and even when such information appears, it is often limited only to a name, with no further biographical information available. We may indeed question why it matters at all to say that a text was written by a particular individual, when we know next to nothing about that individual. Further, authorship in the Middle Ages was inevitably, in one way or another, a collaborative endeavour, with texts manipulated and changed by multiple individuals, often operating centuries apart, as texts were copied and reworked, so it is arguable that the modern concept of a single author for a given text is not applicable anyway.

Medieval conceptions of authorship were also very different from those which we follow today. Many readers in the later centuries of the Middle Ages appear to have been interested in the individual responsible for a text, as evidenced by a shift which occurs in biblical exegesis in the thirteenth century, from a focus on God as the spiritual inspiration for Scripture to a focus on the human *auctor* behind the text (Minnis 1988: 5). Yet when it comes to literary texts which are left unattributed, both in manuscript paratext and within the texts themselves, we may wonder whether the earliest readers of these texts were indeed interested in the identity of their author. Perhaps the attribution of authorship was simply not as important as it is for us today, due both to the inherent complexity of authorship in the period, and also the ways in which society considered and managed ideas surrounding intellectual property.

Nevertheless, we ought to be cautious of suggesting that questions of authorial attribution in the Middle Ages are without merit to modern literary scholarship, and that an 'auteur perdu' should ever remain thus. Besides the obvious human curiosity of wanting to discover the

'maker' behind a work of art, literary or otherwise, there is a value to the attribution of authorship as a method of textual grouping: knowing that two texts share an author may tell us something about that author, but it tells us a great deal more about the texts. In identifying literary texts which share an author, we are alerted to the presence of extra- and inter-textual features which lend unity to a given corpus, revealing the similarities shared by a given group. On the other hand, by asserting that a corpus is the work of one author, we are also making an exclusionary claim, which differentiates a given corpus from other texts, whether or not they have features in common.

These two effects, of unity and distinctiveness, allow for a particularly powerful reading of a given group of medieval texts, not least since we are more likely to read them side-by-side, as their earliest readers may have done. We come to see a single-author corpus as an interconnected network, with the themes of one text calling to mind those of another. Elements of one text which may be lost in isolation become critically important to our interpretation when that text is taken in the context of another where it is more prominent.

In this chapter, I show that digital and statistical approaches have much to offer the examination of uncertain authorship. I take a corpus of texts found mainly within fr24432,[1] and examine the case for its attribution to one Jehan de Saint-Quentin. Through this analysis, we reinforce the attribution of 24 texts to Jehan made by previous scholarship and reach the novel conclusion that there are further works which deserve inclusion in his *œuvre*, specifically *Le dit de Guillaume d'Engleterre* and *Le dit de Robert le Diable*, both of which appear within fr24432.

The main methodological contribution of this chapter, meanwhile, is the suggestion that methods of digital, statistical, and codicological analysis are at their most effective when used

[1] Again, for an overview of the texts in fr24432, including those attributed to Jehan de Saint-Quentin, see Appendix A: fr24432 contents on p. 285.

alongside the linguistic and textual approaches common in previous scholarship, allowing us to enhance or contradict the attribution of texts to a specific author through concrete, reproducible results. As with the previous chapter's contributions to codicology and palaeography, then, this chapter makes specific contributions to the study of individual textual objects, but also adds further tools to the literary scholar's toolbox, demonstrating the utility of its methods to contexts beyond the scope of this chapter, and reshaping the ways we think about authorship in the Middle Ages.

3.1.1 Jehan de Saint-Quentin

In line 221 of *Du chevalier et de l'escuier* (ff. 112ra–115^{rb}), an author-narrator refers to himself as 'Jehan de Saint-Quentin.' This is the only time this name appears in fr24432, and it is not found in any other manuscript from the period, with this text an *unicum* and with no known contemporary references existing elsewhere.

Even assuming that Jehan de Saint-Quentin ever existed, we therefore know nothing at all of his biography, not even whether he was a *clerc*, as some have suggested (Grente *et. al.* 1994: 847), or a *jongleur* (Morawski 1939: 327).[2] The most likely source of his byname, meanwhile, is the town found in the Aisne region, leading Reenen and Schøsler (1997: 595) to take Jehan's language as representative of the variety used across the whole of the region. As we noted in Chapter 1, however, there is no evidence that Jehan spent any time at all in Saint-Quentin, and medieval toponymic bynames are notoriously unreliable as a method of determining a person's place of origin. As for the period in which he was active, Léonard (1996: 248) appears justified in hedging her bets and opting for the 'première moitié du XIV^{e} siècle,'

[2] This is not a discussion which we will pursue in further detail, given its minimal relevance for the question of Jehan's possible authorship of the texts we will examine, but suffice it to say that the boundaries of such categories are in any case often significantly blurred.

given that our best guess at dating *Du chevalier et de l'escuier* comes only from the simple fact that most of the datable texts within fr24432 fall within that period.

Jehan de Saint-Quentin would therefore appear to be an almost wholly unknowable figure, an 'auteur perdu' to the winds of time, to whom only 228 lines of poetry can be attributed with even the slightest degree of certainty. Nevertheless, a small body of scholarship has developed over the course of the last century which suggests that his contribution is more significant, and that Jehan ought to be recognised as the author of a number of other poems within fr24432, as well as of some in other manuscripts.

The case study in this chapter therefore seeks to answer a simple question, namely whether the use of new methods made possible by technological and scholarly developments since the last major work on Jehan's *œuvre* can assist in supporting or refuting the attribution of these texts. In each section of the chapter, I examine some feature of the texts attributed to Jehan, both considering the validity of previous analyses and pursuing new directions. In doing so, I seek not only to bring further clarity to the question of which texts may be confidently attributed to this one author, but also to demonstrate the utility of digital and statistical techniques as a means of answering diverse research questions regarding the attribution of authorship.

3.1.2 Existing attribution to Jehan de Saint-Quentin

Before we delve into our analysis of Jehan's work, it is useful first and foremost to examine the scholarly development of the attributed *œuvre* over time. The first attribution of any text other than *Du chevalier et de l'escuier* to Jehan de Saint-Quentin came when Gustav Gröber, in his *Grundriss der romanischen Philologie* (1888–1902), isolated a group of texts which he believed could be attributed to Jehan. His corpus totalled 20 dits, all found within fr24432, which are

listed in Table 3.12. Gröber was primarily interested in the texts' sources, and gave no evidence or explanation for his inclusion and non-inclusion of any texts.

Table 3.12 – texts attributed to Jehan de Saint-Quentin by Gröber (1888–1902).

Title	Item in fr24432
Un dist que on clamme respon	3
Le dit des trois chanoinnes	11
Le dit des III pommes	20
Le dit de la bourjoise de Romme	28
Le dit des II chevaliers	30
Le dit de l'enfant rosti	31
Le dit du povre chevalier	32
Du chevalier et de l'escuier	33
Le dit de la borjoise de Narbonne	34
Le dit du chevalier qui devint hermite	35
Le dit du cordouanier	36
Le dit du petit Juitel	37
Le dit de l'enfant qui sauva sa mere	38
Le dit de l'eaue beneoite et du vergier	39
Le dit du riche home qui geta le pain a la teste du povre	40
Le dit du chien et du mescreant	41
Le dit de la pecherresse qui estrangla III enfans	42
De Florence de Romme	55
Le dit du buef	61
Le dit de la beguine	74

Following this, Joseph Morawski published an analysis of a corpus of 22 dits which, he claimed, could all be attributed to Jehan de Saint-Quentin, a so-called 'jongleur oublié' (1939: 327). Moraswki's list differed from that of Gröber through the inclusion of an extra dit from within fr24432, *Le dit de Merlin Mellot*, as well as two from outside the manuscript (*D'une abesse que Nostre Dame delivra de confusion* and *La Vie sauveur l'ermite*). He also excluded *De Florence de Romme*, which, by contrast, Gröber had considered worthy of inclusion in his corpus. The list of poems which Morawski attributes to Jehan can be found in Table 3.13.

Unlike Gröber, Morawski was explicit in his exclusion of texts which, he argued, shared notable similarities with the *dits* in his corpus, but were not sufficiently similar to be considered the works of Jehan. In his 1939 work, he explained his exclusion of *De Florence de Romme* and *Le dit des anelés*, noting that 'nous croyons que ces deux *Dits*-romans sont l'œuvre

d'un même auteur, mais nous n'avons pas acquis la conviction qu'ils soient de Jean' (1939: 351–352). Morawski also references a number of other texts which, he believes, are possible contenders for inclusion in Jehan's *œuvre*, but for which insufficient evidence exists to reach a definitive conclusion.[3] In 1940, Morawski returned to the topic of the corpus of Jehan's *dits*, but did not add any further texts, instead focussing on the explicit exclusion of a further four texts from the corpus,[4] and reinforcing his conclusion that *Le dit des anelés* and *De Florence de Romme* are not the works of Jehan, through the use of a largely linguistic approach which we will examine in detail below.

Table 3.13 – texts attributed to Jehan de Saint-Quentin by Morawski (1939: 328–330).

Title	Item in fr24432	Morawski's ref.
Un dist que on clamme respon	3	V
Le dit des trois chanoinnes	11	XX
Le dit des III pommes	20	XIX
Le dit de la bourjoise de Romme	28	VII
Le dit des II chevaliers	30	II
Le dit de l'enfant rosti	31	XIII
Le dit du povre chevalier	32	III
Du chevalier et de l'escuier	33	I
Le dit de la borjoise de Narbonne	34	VIII
Le dit du chevalier qui devint hermite	35	IV
Le dit du cordouanier	36	VI
Le dit du petit Juitel	37	XIV
Le dit de l'enfant qui sauva sa mere	38	XV
Le dit de l'eaue beneoite et du vergier	39	XVI
Le dit du riche home qui geta le pain a la teste du povre	40	XVII
Le dit du chien et du mescreant	41	XVIII
Le dit de la pecherresse qui estrangla III enfans	42	X
Le dit de Merlin Mellot	53	XII
Le dit du buef	61	XXI
Le dit de la beguine	74	XI
La Vie saint Sauveur l'ermite	Not present[5]	XXII
D'une abesse que Nostre Dame delivra de confusion	Not present[6]	IX

[3] Specifically, the alexandrine quatrain versions of the lives of Saint Alexis, Saint Dieudonnée, and Saint Gregory. For editions, see Paris & Pannier (1872), Thornton (1932), and Fant (1887) respectively.
[4] *La Vie saint Basille*, *Le Miracle de saint Servais*, *La Vie de saint Jehan Paulus*, and *Le Jugement Richart*. For an edition of the first text, see Denomy (1956), for the second and third, see Kleist (1973). The fourth text is unedited, and appears in manuscript form only in the early-16th-century Paris, BnF nouv. acq. fr. 10721, ff. 83r–93r, but also appears in an incunable entitled *S'ensuyt le Rommant de Richart filz de Robert le diable qui fut duc de Normandie*, which Morawski dates to around 1496 (1940: 516). Morawski also briefly mentions *Le dit de Robert le Diable*, a text which appears within fr24432, but does not go so far as to say that it is not the work of Jehan. We will return to this text later in our analysis to consider whether it could in fact merit attribution.
[5] Appears only in Paris, BnF Ars. 2115, ff. 48r–v, 52r–57v.
[6] Appears only in Paris, BnF fr. 12483, ff. 66ra–68rb.

Birger Munk Olsen (1963a, 1963b, 1978) took up and examined Morawski's claims regarding Jehan's authorship of the *dits*. He is largely critical of Morawski's approach to attribution, but does not take issue with the majority of his decisions to exclude certain texts. Munk Olsen's work on the *dits* culminates in his 1978 monograph, the first attempt at a complete edition of Jehan's *œuvre*, which contains a total of 24 *dits*, including all of those cited by Morawski, with the addition of *Le dit des anelés* and *De Florence de Romme*, the two poems which Morawski had believed to have been composed by a single author, but which he had been unwilling to accept as being by Jehan. Munk Olsen claims that his own analysis, which we return to in greater detail throughout this chapter, points to the fact that these two texts are sufficiently similar to the rest of the corpus, in both language and style, to warrant re-inclusion. Based on an analysis of a number of factors, he also rejects the possibility that a version of the *Vie de saint Alexis* in alexandrine rhyming quatrains could be a work of Jehan.[7] Munk Olsen's 1978 list can be found in Table 3.14.

Table 3.14 – texts attributed to Jehan de Saint-Quentin by Munk Olsen (1978).

Title	Item in fr24432	Munk Olsen's ref.
Un dist que on clamme respon	3	A
Le dit des trois chanoinnes	11	B
Le dit des III pommes	20	C
Le dit de la bourjoise de Romme	28	D
Le dit des II chevaliers	30	E
Le dit de l'enfant rosti	31	F
Le dit du povre chevalier	32	G
Du chevalier et de l'escuier	33	H
Le dit de la borjoise de Narbonne	34	I
Le dit du chevalier qui devint hermite	35	J
Le dit du cordouanier	36	K
Le dit du petit Juitel	37	L
Le dit de l'enfant qui sauva sa mere	38	M
Le dit de l'eaue beneoite et du vergier	39	N
Le dit du riche home qui geta le pain a la teste du povre	40	O
Le dit du chien et du mescreant	41	P
Le dit de la pecherresse qui estrangla III enfans	42	Q
Le dit de Merlin Mellot	53	R
De Florence de Romme	55	S
Le dit des anelés	57	T

[7] The so-called *A4* version, edited in Paris & Pannier (1872).

Le dit du buef	61	U
Le dit de la beguine	74	V
La Vie saint Sauveur l'ermite	Not present	W
D'une abesse que Nostre Dame delivra de confusion	Not present	X

Since this work, very little attention has been paid to the attribution of these poems, with the (usually very cursory) references to Jehan's *œuvre* generally taking Munk Olsen's corpus as the definitive list.[8] The only more recent attempt to engage with the question of attribution comes in a footnote to a recent article, in which Gabriel Cholette (2022: 96, n. 6) questions why *Le dit de Robert le Diable* is not attributed to Jehan, but does not provide evidence to support a shared attribution beyond the fact that the text is also found in fr24432, and 'ressemble beaucoup aux autres dits.' He makes no mention of the attribution of *Le dit de Guillaume d'Engleterre*, and argues that the similarities in 'la versification, le style, la fonction du récit [...] relève[nt] davantage du genre littéraire que de l'auctoralité' (2022: 96), a claim which we will dispute on the basis of comparisons to other single-genre corpora. He is however correct to suggest that 'le fait que les dits narratifs aient été regroupés dans un même manuscrit n'est pas suffisant pour affirmer qu'ils ont été écrits par le même auteur.' While collective compilation is a mark in favour of the shared authorship hypothesis, we must go much further in order to be able to say with any degree of certainty that these poems ought to be described as the work of a single author.

3.2 Methodology

Progress in recent decades on digital and statistical approaches to the analysis of texts allows us to approach Jehan's *œuvre* from directions unavailable to previous scholarship. So many of Morawski and Munk Olsen's arguments in particular are founded on quantitative claims, and yet as we shall see throughout our analysis, their data nevertheless lack the reliability and reproducibility that more modern approaches can offer. The aim of this chapter is therefore

[8] See, for example, Türk (2000).

to build on previous examination of Jehan's *dits* through the use of digital and statistical approaches in particular, presenting methods that can be applied to other research questions within the study of medieval literature and beyond. In some cases the arguments which I present are intended to be complementary to earlier arguments, building on scholars' work using methods that were unavailable to them, in order to enhance and strengthen the conclusions they reach. In others, particularly in cases where scholarly thought has moved on since the most recent work on Jehan's poetry, I offer alternative approaches to the examination of the corpus to those which have been attempted previously.

While many of the methods in this chapter overlap with what has come to be known as stylometry, the aim here is not simply to apply existing methods of analysis to a new context. Rather, the intention is to show the value of taking a bespoke, text-focussed, and data-driven approach to textual analysis, through the application of methods developed to suit the specificities of the corpora which we are to examine.[9]

3.2.1 Use of statistical methods

Following standard practice, when statistical correlation is calculated for a given feature, it is described as strong when the correlation coefficient is between ± 0.5 and 1, and weak when it is between ± 0.25 and 0.5. Where statistical data have been extracted which allow us to contrast use of a given feature, p-value tests have also been conducted in order to determine whether a distinction between two sets of data can be said to be statistically significant, rather than simply a function of random sampling. In cases where we compare the use of a given feature in one corpus to that in another, the observed and expected values are often too small to allow, for example, for the use of a Chi-squared 'goodness of fit' test. In such cases, some

[9] Eder *et al.* (2016) is one good example of the existing tools that can be applied to questions of digital stylometry, but which we have not made use of here.

(such as Campbell 2007) have recommended a combined method involving the Chi-squared and Fisher–Irwin tests, but in the interests of uniformity I have ultimately chosen where possible to make consistent use of a randomisation method, run across a million samples, which results in a reliable p-value regardless of the number of observed values.[10] The randomised nature of this test means that the p-value will always fluctuate somewhat, but not enough to vary the result of the test. Where necessary, p-values are given to five decimal places, and the threshold for statistical significance is set at $p < 0.05$. When comparison is made between Jehan's texts (the 'target' corpus) and some reference corpus, the null hypothesis in all cases is that the rate of occurrence of an observed feature of the target corpus does not differ from that of the same feature in the reference corpus.

3.2.2 Target corpus

As the central object of study for this chapter, I take the texts contained within Munk Olsen's (1978) corpus of Jehan's *dits*, a corpus of some 63,347 words which I term the *J-corpus*. This corpus includes both the texts attributed to Jehan within fr24432, as well as the two texts found outside it. While we will critically examine the boundaries of this corpus, it is nevertheless a good starting point since it is this corpus of texts which has been broadly accepted by current scholarship as being the work of Jehan.

In order to facilitate the extraction of reliable and reproducible data from the J-corpus, it has been necessary to make use of the P5-compliant TEI-XML encoding completed at the outset of this project, whose encoding principles are discussed in the Introduction. When encoding the texts of the J-corpus, I have paid particular attention to Munk Olsen's (1978) edition of the poems he attributes to Jehan, and for these texts I have gone beyond the standard encoding principles used elsewhere in the codex to mark direct speech using the `<q>` tag.

[10] The R script for this calculation can be found in Script 3.1 on p. 317.

In instances where the analysis requires us to search directly within the encoded content, such as in our examination from page 229 on the moment of the onset of direct speech, the required data have been extracted directly from the XML encoding through the use of XPath and Regular Expression queries. In most cases for the analysis in this chapter, however, the required data have been more textual in nature, and as such it has been necessary to produce a version more akin to a single-manuscript critical edition, so that analysis can be carried out on a more standardised version of the text. This has been completed through an XSLT transformation, which expands abbreviation and accepts appropriate additions, deletions, regularisations, and corrections within the encoding. The resulting text file somewhat resembles a critical edition but should not be interpreted as anything other than a source of textual data, since it does not engage with questions of textual variance, where a text appears outside of fr24432, and it does not contain modernised punctuation, with the exception of quotation marks, so would be unsuitable for human reading. In the case of those texts which are not found within fr24432, *Sauveur* and *Abbesse*, Munk Olsen's (1978) edition has sufficed, with data that are extracted from the XML encoding for other texts in this case simply extracted manually, given the brevity of both poems.

Digital and textual analysis of the corpus has proceeded from these text files in two separate directions. In the first, the text files themselves form the data from which analysis can proceed. This is the case for the analysis of word forms within the corpus which do not vary in spelling or form, as in the examination of the use or non-use of *mon*, *ton*, and *son* as possessive adjectives preceding a feminine vowel-initial noun, for which the orthographic forms do not vary, and which we shall examine in detail from page 245. Analysis of this type has largely been performed using Antconc (Anthony 2023).

Most examination of the corpus, however, has followed a second path from the encoding to eventual analysis which involves a further step of processing before the data can be usefully

applied. In instances where word forms vary either in inflection or in spelling throughout the corpus (in most cases both are true), they have been lemmatised. That is, we assign an arbitrarily defined 'dictionary' form to each instance of a given lexeme, to extract clearer data on the use of a particular lexeme within the noise generated by scribal and inflectional variation.

Specifically trained models exist for this purpose, including RNNTagger (Schmid 2019), a part-of-speech tagger and lemmatiser, which uses machine-learning techniques to reach probabilistic assertions regarding both the part of speech and associated lemma of any given word form in a text.[11] Given the inherent variability in the Old French orthography, as well as the fact that RNNTagger is a probabilistic model trained on manually tagged text, it is perhaps unsurprising that the output is not always perfect. Table 3.15, for example, contains the output of the first stanza of *Abesse*, once it has been passed through RNNTagger. Issues with the tagging include that 'dit' has been tagged as a finite verb rather than a noun, while the two forms 'oïr' and 'orront,' are given two different lemmata: *öir* and *ouïr*, despite being forms of the same verb, likely due to discrepancies in the training data.

An alternative approach would be to run the data through a more generalist AI natural language model, such as ChatGPT (OpenAI 2022a). Table 3.16 contains a version of the same stanza tagged and lemmatised using this method. Advantages of this model include the fact that specific details of the word forms, such as verb tense, are included in the tagging, and the model is more successful at providing consistent lemmata across different forms. This tagging was only achieved after some human feedback, however, since the initial output was

[11] In addition to RNNTagger, see for example its precursor TreeTagger (Schmid 1994), which uses more straightforwardly probabilistic rather than machine learning methods. The 2020 *Dictionnaire du Moyen Français* (DMF; ATILF 2020) also contains a lemmatiser, but it is unsuited to the task at hand, since it shows only a lexicographic, and not a syntactic, understanding of a given text, and offers all lemmata which match a given form, rather than selecting the most likely one.

less successful, mistakenly claiming, for example, that 'voudront' is conditional, and citing 'm'orront' as a single word form.[12]

In the case of the first major error, a model trained on manually tagged text would likely not have a single example of 'voudront' tagged as conditional, unless there were some mistake in the training data, and so such errors would be unlikely to occur. ChatGPT, meanwhile, for which the training data are far more generalised, has *hallucinated*, that is it has given the response that it predicts is most likely to be accepted, despite the fact that it cannot in any way be 'true.' The second major error, ChatGPT's inability to recognise 'm'orront' as two words, is a simple error of segmentation that suggests an underlying inability to separate words, again since the model has received no direct training in this respect. This is of course concerning for a model that purports to be able to carry out part-of-speech tagging, and points towards the possibility of similar segmentation errors occurring elsewhere.

Table 3.15 – vv. 1–4 of Abesse, *tagged and lemmatised by RNNTagger.*

Word	POS Tag	Lemma
Ceus	PROdem	cil
qui	PROrel	que
voudront	VERcjg	vouloir
oïr	VERinf	öir
un	DETndf	un
biau	ADJqua	bel
dit	VERcjg	dire
recorder	VERinf	recorder
Et	CONcoo	et
qui	PROrel	que
a	PRE	a
Nostre	DETpos	notre
Dame	NOMcom	dame
se	PROper	soi
voudront	VERcjg	vouloir
acorder	VERinf	accorder
,	PONfbl	,
Si	ADVgen	si

Table 3.16 – vv. 1–4 of Abesse, *tagged and lemmatised by ChatGPT.*

Word	POS Tag	Lemma
Ceus	PRO:DEM.M.PL	ceux
qui	PRO:REL	qui
voudront	VER:futur	vouloir
oïr	VER:infi	ouïr
un	DET:M:S	un
biau	ADJ:M:S	beau
dit	N:M	dit
recorder	VER:infi	recorder
Et	CC	et
qui	PRO:REL	qui
a	PRP	à
Nostre	DET:F:S	notre
Dame	N:F	dame
se	PRO:PER	se
voudront	VER:futur	vouloir
acorder	VER:infi	accorder
	—	
Si	ADV	si

[12] We could also note the fact that ChatGPT appears not to be able to treat sentence-internal punctuation as tokens, although given that I have not standardised the punctuation in these texts, this is of minimal importance here.

se	PROper	soi	se	PRO:PER	se
traient	VERcjg	traire	traient	VER:pres_subj	traire
vers	PRE	vers	vers	PRP	vers
moy	PROper	moi	moy	PRO:PER	moi
,	PONfbl	,		—	
si	ADVgen	si	si	ADV	si
m’	PROper	je	m	PRO:PER	me
orront	VERcjg	ouïr	orront	VER:futur	ouïr
recorder	VERinf	recorder	recorder	VER:infi	recorder
Comment	ADVsub	comment	Comment	ADV	comment
la	DETdef	le	la	DET:F:S	le
mere	NOMcom	mere	mere	ADJ:F:S	mère
Dieu	NOMpro	Dieu	Dieu	N:M	dieu
scet	VERcjg	savoir	scet	VER:pres	savoir
les	DETdef	le	les	DET:M:PL	le
bons	NOMcom	bon	bons	ADJ:M:PL	bon
conforter	VERinf	conforter	conforter	VER:infi	conforter
.	PONfrt	.	.	SENT	.

Both errors therefore reveal the significant deficiencies of ChatGPT for the task of lemmatisation for Old French in that, as a natural language model, it has not been trained specifically for the purposes of lemmatisation, and so does not necessarily have any underlying conception of the linguistic content, rather it simply produces an output which the model predicts will be accepted as correct (OpenAI 2022b).[13] Given the presence of these errors, as well as the difficulties in obtaining reproducible outputs from ChatGPT at scale, I have therefore ultimately preferred the output of RNNTagger, on the understanding that the tagging may not always be completely accurate.[14] Approaches based on word forms and lemmata are thus both used side by side to ensure that relevant appearances of words are not missed, in order to achieve the most accurate results possible. Analysis using these data has primarily been performed using Sketch Engine (Lexical Computing 2022).

[13] The version of ChatGPT used here is version 3.5, released early 2023. Later versions may be more successful in certain respects, but will still suffer from the same difficulties regarding limited prompt and output length, as well as the reproducibility of results.

[14] RNNTagger attempts to provide morphosyntactic tagging for some languages, such as Modern French, but does not do so for Old French. This is not a problem for this dataset, since such information is not necessary to our analysis, but ChatGPT’s attempts to also tag morphosyntactic features of the source text (with the notable exception of case) mean that for other projects where such tagging is necessary, it could be a preferable tool.

3.2.3 Reference corpora

When making claims about whether a particular feature of a text is 'normal' or somehow 'unusual,' it is easy for such arguments to be based solely on a scholar's sense of what qualifies for such descriptions, as is occasionally observed in the existing literature on Jehan's work. To allow for more evidence-based arguments, it is necessary to provide reference corpora which our hypotheses may be tested against. The analysis in this chapter makes use of five reference corpora, each of which serves a specific comparative purpose.[15]

The *P-corpus* (279,204 words) consists of the 64 texts found within fr24432 which are not part of the J-corpus, that is, they have not been previously attributed to Jehan.[16] The contents vary in genre, form, style, chronology, and geographic origin, and so the corpus is not intended as a like-for-like comparator for the J-corpus. Instead, it allows comparison between the texts of the J-corpus and those which, despite varied origins, came to appear in the same manuscript as most of the J-corpus. It is this corpus which therefore allows us to look at fr24432 as its earliest readers may have done, to take a holistic approach to the codex in order to compare similarities and differences between different textual groupings in the same manuscript.

Given the significant variation present throughout the texts of fr24432, it has sometimes been necessary to subdivide the P-corpus to provide the appropriate data for the method of analysis at hand. The broadest subdivision of the corpus, which I term the *P0-corpus* (278,679 words), eliminates only the two highly fragmentary poetic texts present in the P-corpus which could threaten to skew the data when our analysis concerns the proportion of a text that displays a certain feature, given that only part of the texts survive.[17] The *P1-corpus* (160,741

[15] These corpora have all been lemmatised with RNNTagger to allow word-based and lemma-based analyses.
[16] The name of the corpus corresponds to the siglum which is often assigned to the whole of fr24432, in particular in descriptions of the manuscripts containing *fabliaux* (cf. Noomen & van den Boogaard 1983–1998).
[17] *Lai d'amour 1*, and *Trois mortes*. *Patenostres*, a third fragmentary poem, is sufficiently long to still merit inclusion, at 275 lines, since it appears that only a few lines are missing at the beginning of the text.

words), meanwhile, includes only those texts from the P0-corpus which are consistent in that they are all entirely in verse, and again excludes the fragmentary texts. This sub-corpus contains some 55 texts in a variety of metres and forms, and is on occasion therefore subdivided again to include only the 7 texts which appear in alexandrine monorhyme quatrains,[18] the poetic form found throughout the J-corpus (the *P1a-corpus*, 31,813 words). Within this P1a-corpus I include *Guillaume* and *Robert*, the two texts from within fr24432 which we will later propose as additional works by Jehan. Where demanded by the analysis, these two poems are frequently considered apart from their containing corpora.

Our second reference corpus, which I term the *B-corpus* (6,382,676 words), is by far the largest, consisting of the 219 texts contained in the 2022 release of the *Base de Français Médiéval* (BFM; Guillot-Barbance *et al.* 2022).[19] This corpus contains a broad array of texts, ranging from the ninth to the fifteenth centuries across several dialect areas and in a mixture of genres, in both verse and prose. Therefore, as with the P-corpus, the B-corpus would be an inappropriate source of data to use as a direct comparator for the J-corpus, in that differences between the J- and B-corpora can often likely be explained by the temporal and generic variety of the latter corpus, rather than confirming the single-author hypothesis for the J-corpus. However, the B-corpus does have utility in that it allows the comparison of the J-corpus with a much broader spread of medieval French texts, and can help us identify the extent to which trends observed in the target corpus may have been present more broadly in written forms of medieval French. Subsets of the B-corpus, including those texts with closer

[18] *Guillaume*, *Bible*, *Vision*, *Vie du monde*, *Mais*, *Robert*, and *Menage*.
[19] RNNTagger was in fact trained on several texts within the B-corpus which have been manually lemmatised (Lavrentiev 2023). Rather than use only the manually tagged texts, however, I have chosen to re-tag the whole corpus from plain text, so as to match the tagging of the other corpora as far as possible, given that any tagging errors are more likely to be consistent across all the corpora if they have been tagged in the same way.

formal, generic, and temporal connections to the J-corpus, have been consulted where necessary for a particular direction of analysis.

The *F-corpus* (28,735 words) contains texts from an established generic grouping, consisting of 22 *fabliaux* transcribed and encoded as part of a forthcoming more comprehensive data release by the Laboratoire IHRIM at the ENS de Lyon (Pierreville, 2023).[20] The encodings were created by the same team who produced the texts of the *Base de Français Médiéval* used within the B-corpus, and thus the transcription methodologies of the two corpora are comparable. The texts of the F-corpus are the work of a number of different authors, therefore allowing us to compare the texts attributed to Jehan to those which fit comfortably within the same generic grouping, somewhat parallel to that of the J-corpus, but which do not share the same author. We are thus able to identify features of the J-corpus which may be present only because the texts within it are all of the same genre, as well as instances when these features can be said to be more comfortably the result of shared authorship.

The *L-corpus* (33,443 words), meanwhile, offers a counterpoint to the F-corpus. It contains the 13 *Lais* of Marie de France (including the *Prologue*) as presented in Warnke's 1900 edition (this being the most recent edition available in a digital format), and thus is largely comparable to the J-corpus in that its texts are consistent in genre and form, all while being said to all be the works of the same author. When used in conjunction with the F-corpus, it allows us to identify which features of the J-corpus may result from the corpus' generic conformity, and which may result from shared authorship.[21]

[20] For details of the texts included in this corpus, and for the manuscripts from which they are drawn, see Appendix 3A: F-corpus contents on p. 311. 19 of these texts are edited in Dufournet (1998), and they were chosen by the IHRIM team for pedagogical reasons, Dufournet's edition being a set text for the 2024 *Agrégation externe de lettres modernes*. Despite this somewhat arbitrary sampling process, the texts nevertheless present a sufficiently broad base to be useful for our purposes here.

[21] Of course, Marie's authorship of the *lais* and her identity are far from undisputed, but the corpus is appropriate for our purposes as it remains likely that the *lais* do share an author, whoever that may have been.

Our final reference corpus allows us to take this analysis even further. The *R-corpus* (77,108 words) consists of the complete works of Rutebeuf, as contained within Bastin & Faral's edition (1959–1960), and accessed via the *Rutebeuf.be* project (Sentjens 2023). Rutebeuf's *œuvre* contains a great deal more generic, formal, and stylistic variety compared to the texts contained within the L-corpus, and therefore the R-corpus allows us to examine the J-corpus from a further critical dimension, considering which of its features may be explicable by the shared author hypothesis, as opposed to features which may simply be explained by the fact that the texts of the J-corpus share formal and generic similarities.[22]

These five reference corpora and their subdivisions therefore give us the opportunity to consider the J-corpus from a number of new perspectives. Rather than imposing a modern sense of what counts as a suitable criterion for shared authorship, the methods contained within this chapter will allow us to encounter the texts of the corpus in amongst the enormous diversity of medieval French literature, without seeking to simplify it, just as the texts' first readers likely did. In so doing, I present an approach that is widely applicable to other questions of uncertain medieval authorship.

3.3 Form

An obvious place to begin our examination of the texts attributed to Jehan de Saint-Quentin is with *Chevalier et escuier*,[23] given that this is the only poem which appears to contain an act of self-attribution by the author himself. This text is a short narrative poem, telling of a once-rich knight on *errance*, who becomes companions with a poor squire. They encounter the Devil, who offers them great riches if they become his followers. After they both deny God and all the saints, the knight refuses to take the final step and deny Mary, so the companions

[22] The F-, L- and R-corpora all contain texts found also within the P-corpus. In all cases, however, the copies of these texts in the P-corpus have not provided the model for the transcriptions of the texts in the other corpora.
[23] From this point on, I return to referring to all texts in the J- and P-corpora by their short titles. Full titles can be found in Appendix A: fr24432 contents on p. 285.

part ways. The knight goes to pray before an image of Mary, who is seen miraculously comforting him, and so a local lord gives him a title, wealth, and his daughter in marriage. The squire, meanwhile, is abandoned to the hangman's noose by the Devil, and the narrator warns sinners of taking the Devil's path. In the second to last stanza of this epilogue, the narrator says of the Blessed Virgin:

> Jehan de Saint-Quentin dit que l'en doit servir :
> On peut de son service a grant honor venir.
> C'est le tresor de grace por tous ceulz garantir
> Qui la veillent de cuer et amer et chierir. (*Chevalier et escuier*, vv. 221–224)

The use of the third person here means that there is not an immediate one-to-one link between Jehan and the *je* of the narrator. It is not unreasonable, however, to assume that this is indeed an act of self-attribution, given the extent to which the claims of this 'Jehan de Saint-Quentin' overlap with the didactic intent of the narrator, and previous work on other authors has established that third-person self-reference by an author is not an uncommon phenomenon.[24] We can therefore be reasonably secure in the attribution of this text to Jehan.

The later collective attribution of the rest of the J-corpus to Jehan is made possible by the fact that *Chevalier et escuier* is an excellent archetype of the other texts for a number of reasons, not least given that, at just 228 lines of verse, it is short. It fits comfortably within the range of lengths found in the J-corpus, from 152 metrical lines (*Cordouanier* and *Pecherresse*) to 832 (*Florence*).[25] The text also begins with a four-line prologue, and ends with an eleven-line epilogue, features which Morawski (1939: 330) notes are present across all the *dits* he attributes to Jehan, and which we likewise find in the texts added by subsequent scholarship.

[24] See, for example, Marnette (1998).

[25] The length of *Florence* supplied here does not exactly correlate with the length of the text as it appears in fr24432, given the absence of folio 223 from the codex, which I assume to be a later loss. Comparison with Grenoble, Bib. mun., 871, the other manuscript in which *Florence* appears (on which, see the codicological appendix on p. 318), confirms that lacuna consists of 72 lines of verse, corresponding to a single folio containing four 36-line columns, with each alexandrine line split into two hexasyllable hemistichs.

Morawski (1939: 330) observes that these prologues and epilogues in the corpus share a number of similarities of content. The prologues, he notes, contain exhortations to a putative audience, who are called to listen and learn from the text's content. In this sense, the prologue to *Chevalier et escuier* once again offers a representative example of those found elsewhere:

> Moult est bone acointance d'estoire, bone gent,
> Bien m'en puis percevoir et en maint lieu souvent
> Que quant je dis un dit a tous generaument,
> Car il nous couvient vivre des bons premierement. (*Chevalier et escuier*, vv. 1–4)

As for the epilogues, as in that of *Chevalier et escuier*, they often contain praise of, and prayers to, various religious figures, in particular Mary, with exhortative structures such as *or prions* and *si prions* particularly frequent.[26]

Nevertheless, as Munk Olsen (1963b: 151) notes, there is nothing unusual or surprising about the use of prologues and epilogues in Jehan's *dits*, since such sections are commonly found in other texts with the same form. Nor, in fact, do the details of their content set the J-corpus apart. For Munk Olsen, it is the fact that they are consistently present, and are largely consistent in what they discuss, which provides a unity to the J-corpus that supports, if not doing enough to confirm, a collective attribution.

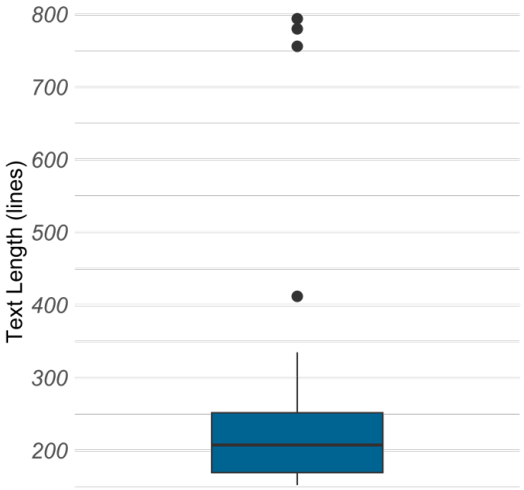
3.3.1 Text length

A more data-driven approach can shed further light on the matter. Building on the findings of this previous scholarship, even simple statistical analysis of the length of the texts in the J-corpus can yield interesting results. Looking at the overall lengths of the texts in the corpus, as represented in the box-and-whisker plot in Figure 3.49, we find that the poems have a median length of 208 lines, and 18 of the 24 texts (75%) come in at under 300 lines in length.[27] The corpus is therefore mostly united in its brevity, but the plot also reveals the

[26] For more on this feature, see the section on Performance on p. 176.
[27] For the data regarding the line, prologue, and epilogue lengths of the J-corpus, see Table 3.47 on p. 312.

significant gap in length between three outlying, longer texts (*Florence*, *Anelés*, and *Buef*), and the next longest text (*Chanoinnes*), which is itself still an outlier despite being nearly half the length of the longest at only 412 lines. These three longer texts all appear relatively close together in fr24432, giving some indication that they were considered, by the compiler of this codex at least, as being a group *à part*, somehow distinct from the other poems in alexandrine monorhyme quatrains found elsewhere in the manuscript.[28]

Figure 3.49 – box-and-whisker plot of the lengths of the J-corpus.

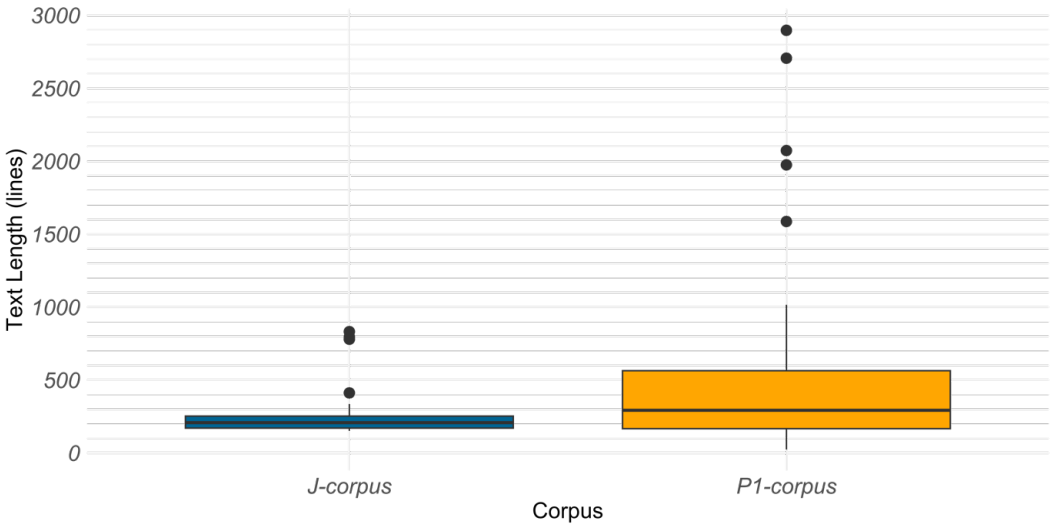


This statistical evidence, combined with the codicological evidence, suggests that we may see Jehan's *œuvre* as dividing into two neat groups, those around the 200-line mark which are written with brevity and concision in mind, perhaps for brief recitation, and those around 800 lines in length, which provide space for a more expansive narrative. Certainly, this does nothing to suggest that the longer poems were not composed by Jehan. If anything, the difference in form makes the similarities in areas other than form, which we will examine throughout this chapter, all the more notable, since they stand out as features that bring a unity to texts which are in some sense disparate.

[28] *Florence* appears on ff. 215^{rb}–226va, *Anelés* on ff. 231ra–241^{rb}, and *Buef* on ff. 247ra–257va. For further discussion of this grouping, see the codicological appendix on p. 318.

Fruitful examination of text length comes also from comparison of the J-corpus to those poems found elsewhere in fr24432, contained in the P1-corpus, which can help us identify how usual the length of Jehan's texts is for the codex as a whole.[29] As presented in Figure 3.50, we observe that the contents of the J-corpus, with a median length of 208, are not unusually short compared to those poems found elsewhere, where the median is 293, but they are certainly on the shorter end of the scale.[30] Furthermore, the longest texts of the P1-corpus, of which the most extreme example is *Tornoiement* at a length of 2,897 lines, reveal that *Florence*, *Anelés*, and *Buef*, while long by comparison to the other texts of the J-corpus, do not appear especially long within the codex as a whole, reinforcing our suggestion that their length does nothing to generate an argument for the texts' exclusion from Jehan's corpus.

Figure 3.50 – box-and-whisker plot of the lengths of the texts of the J-corpus, compared to those of the P1-corpus.



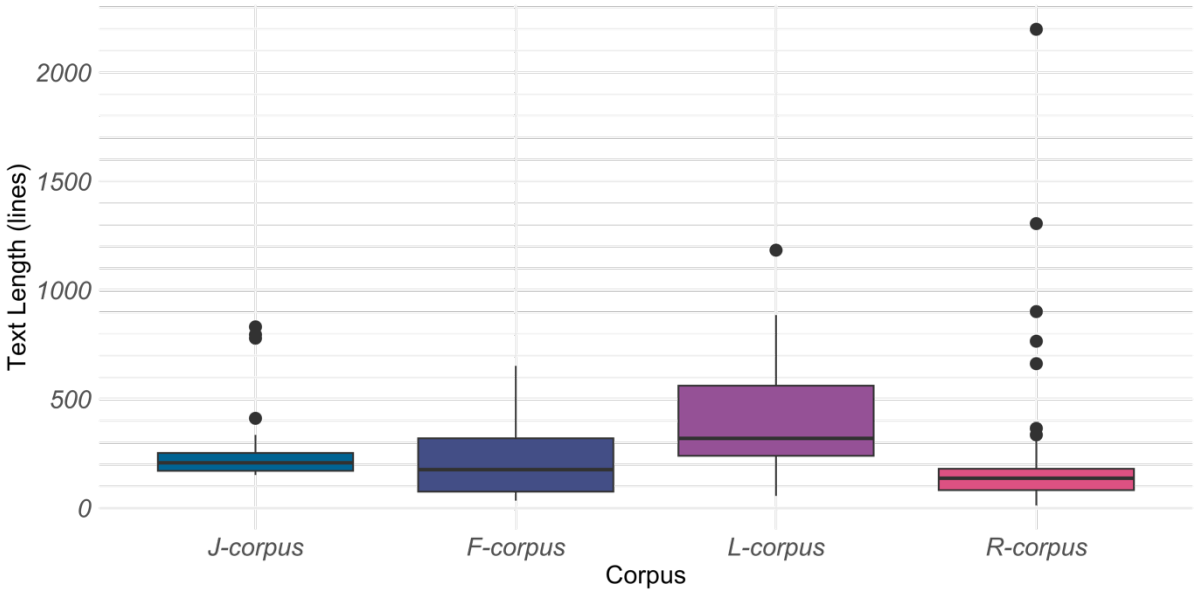
As discussed in our presentation of the reference corpora, however, while it is useful for contextualising the J-corpus in the manuscript in which most of its texts appear, the P1-corpus does not allow like-for-like comparison with the texts of the J-corpus. To do so, we

[29] Note that the P1-corpus excludes three fragmentary poems (*Lai d'amour 1*, *Trois mortes*, and *Vin*).
[30] I should note that my discussion here is of lines of verse for a given text, not orthographic lines of text as they appear in the manuscript. In fr24432, alexandrine verse is usually broken into two hemistichs by the scribes, whereas hexasyllable, octosyllable, and decasyllable are usually presented on a single line. In *Serventois*, no attempt is made for the separations of lines on the page to match the lines of verse, with the division between verse lines instead marked by a *punctus* or similar.

must consider the texts of our more cohesive reference corpora, namely the F-corpus of *fabliaux*, the L-corpus of Marie de France's *lais*, and the R-corpus of Rutebeuf's works.

Figure 3.51 displays box-and-whisker plots for the J-corpus alongside these other corpora, which allows us to make a number of valuable observations. Firstly, while the median text lengths for the corpora are roughly comparable (F: 176.5, L: 320, R: 137), the J-corpus has the narrowest range between the first and third quartiles (J: 82, F: 244.75, L: 322, R: 98), suggesting a level of consistency which exceeds even that of the L- and R-corpora for which the attribution of authorship is more assured, reinforcing the sense of formal unity in the J-corpus which allows us to be more assured in our attribution of these texts to Jehan.

Figure 3.51 – box-and-whisker plot of the lengths of the texts of the J-corpus, compared to those of the F-, L-, and R-corpora.



Comparing the F-corpus to the L- and R-corpora, we also observe that there are no outliers in the first corpus, despite many of the texts being attributed to different authors, whereas an outlier appears in the L-corpus (*Eliduc*, 1,184 lines), and several appear in the R-corpus, of which the most extreme is *La Vie sainte Elysabel*, at 2,198 lines. It is therefore certainly possible for a corpus of texts by the same author to contain texts of lengths which differ significantly from the remainder of the corpus, just as it is possible for the lengths of texts in a corpus

containing a single genre but the work of several authors to be more consistent. As a result, it becomes even clearer that the outliers in the J-corpus are not evidence against a shared attribution, rather they indicate that Jehan, like Marie de France and Rutebeuf, was not a one-trick pony, and was capable of modifying the form of his poems where necessary.

3.3.2 Prologues and epilogues

Taking a similarly statistical approach to the prologues and epilogues of the J-corpus we observe, for example, that *Chevalier et escuier* is somewhat unusual amongst the corpus in that its epilogue is longer than its prologue, a feature not seen anywhere else in the texts of the J-corpus.[31] We could argue that this distinguishes *Chevalier et escuier*, the text in which Jehan signs his name, from those in which he does not. On the contrary, however, it appears that it is precisely because of this act of self-attribution that the prologue reaches this length, rather than suggesting that the whole process of attribution is mis-aligned. The intensity of the epilogue's devotional content suggests that Jehan extended this section, in which he is named, as a means of highlighting his own piety, in a way that he did not do elsewhere.

Examining the P1a-corpus, the 7 texts found within fr24432 which are written in monorhyme alexandrine quatrains and which have not been attributed to Jehan by previous scholarship, we find that unlike the texts of the J-corpus, where each text contains a prologue and an epilogue, this is not always the case in texts of the same form, since fr24432's copy of Rutebeuf's *Vie du Monde* has a four-line epilogue but no prologue.[32]

Turning our attention to prologue lengths, we note that these are usually short in the J-corpus, with the two longest prologues (84 and 40 lines) appearing in two of the longest

[31] *Juitel* has a prologue and epilogue of eight lines each. On the use of narrative approaches to identify the point at which shifts occur between the prologues, narrative content, and epilogues of these texts, see Cholette (2022).
[32] The case of *Guillaume* will be discussed in further detail below. *Menage* is a difficult case, since the voice of the narrator is also the voice of the speaker who begins and ends the text, but the prologue and epilogue in this text arguably consist of just two lines and one line respectively.

texts: *Anelés* and *Buef* (780 and 796 lines), both of which use the prologues to offer a brief summary of the plot to come. Despite some notable anomalies, such as the prologue of 32 lines found in *Beguine*, a text of only 204 lines in length, it is not enormously surprising to find that across the J-corpus there is a statistically significant link between total text length and prologue length ($p = 0.00207$). The correlation coefficient between text length and prologue length is also strongly positive (0.59706), such that we may say that on average longer texts in the corpus are likely to contain longer prologues, a fact which can perhaps be explained through the presence of greater pre-narration of their more complex plots. This observation is also applicable to the P1a-corpus, which produces a correlation coefficient between text length and prologue length of 0.83404 ($p = 0.01966$), suggesting that a strong positive correlation between text length and prologue length is not a feature unique to the J-corpus.

Indeed, we should note a further similarity between the two corpora, namely that for both the J-corpus and the P1a-corpus, the standard deviation of the lengths of the prologues (J: σ = 16.2, 93.3% of the mean; P1a: σ = 12.5, 127.0%), and the texts themselves (J: σ = 204, 70.7% of the mean; P1a: σ = 334.4, 63.8%) are not consistent in the proportion of the mean that they represent, and so the relationship between these two factors will not be linear. In other words, in both corpora an increase in the length of a text results on average in a longer prologue, both in terms of the number of lines, and its relative length as a proportion of the whole text. While this is not a unique feature of the J-corpus, it is nevertheless still an interesting observation that perhaps reflects an increased desire in longer poems in alexandrine quatrains to engross an audience in the story before beginning the narrative, in a way that is not necessary in shorter texts. This hypothesis would also go some way to explain why the longer prologues in the J-corpus in particular often have a significant focus on narrative content, rather than dwelling on any didactic message, since this enables the author to better prepare an audience for what is to come.

A further indicator that trends followed by the prologues and epilogues in the J-corpus are as much determined by formal, generic considerations as they are by authorial ones comes from the fact that for both corpora there is no statistically significant link between total text length and epilogue length (J: $p = 0.2608$; P1a: $p = 0.1595$), or prologue length and epilogue length (J: $p = 0.568$; P1a: $p = 0.2092$). This means that we cannot predict the length of the texts' epilogues based on any other length feature. This absence of a statistical link therefore allows for observations external to questions of authorship, in that it suggests that the content of the epilogue in texts written in monorhyme alexandrine quatrains may, as a norm, have been determined by the immediate requirements of the end of the narrative, be they metrical or thematic, whereas the prologue was determined by the requirements of the text as a whole.

The fact that the standard deviation amongst the epilogues is lower in relative terms than that of the prologues or the texts themselves (J: $\sigma = 3.50$, 54.2% of the mean; P1a: $\sigma = 1.77$, 62.0%) is also a possible indication that authors of such poems saw little reason to vary their epilogues at all. We therefore see something of a cognitive distinction for the authors of these texts between their prologues and epilogues, and thus through our use of statistical methods, we are able to support Morawski's (1939: 330) claim that the epilogues of Jehan's texts conform to reasonably consistent structures, regardless of features of the text to which it is appended, while adding the caveat that this is not a unique feature for the J-corpus.

Much of our examination of the formal features of prologues and epilogues therefore does not mark out the J-corpus in the context of other comparable poems (the P1a-corpus), which circulated in a similar material context to Jehan's texts. However, there is still value in the inclusion of these data, since it suggests a certain formal, and perhaps generic, consistency within Jehan's texts, which contradicts the assumptions of previous scholarship that Jehan's prologues and epilogues quantitatively distinguish his texts from others. While we are thus unable to rely on these data as evidence of the corpus' shared authorship, they do allow us to

better understand the generic framework in which Jehan operated, a question which we will return to in the next section. As we move through the chapter as a whole, the holistic combination of various such unifying features will allow us to reach more assured conclusions regarding Jehan's authorship of the J-corpus and the other texts which we will assign to him.

3.4 Genre

Returning to *Chevalier et escuier*, the text in which Jehan names himself, we may note a further feature which links it to the other texts of the J-corpus, namely its didactic and parenetic function. This is particularly noticeable in the prologue-epilogue pairs, but is found throughout the narrative as well, and it appears to be a necessary characteristic of all the attributed poems, which seek to educate and persuade an audience on questions of good moral behaviour. Morawski notes Jehan's preference for 'des sujets étranges ou horribles, propres à faire hérisser les cheveux sur la tête' (1939: 345–346), which again, as Munk Olsen (1963b: 152) notes, is in no way astonishing for the didactic material of the period, yet it does add a further sense of unity to the corpus, increasing our sense that they are similar enough to share an author. In the case of *Chevalier et escuier*, while its contents may not be as shocking as the infanticide of *Romme* and *Pecherresse*, the two companions' initial willingness to renounce God and the saints, as well as the squire's gruesome ending, is certainly intended to shock, while the knight's redemption provides a contrasting example of good moral behaviour.

When examining the didactic nature of Jehan's texts, previous scholarship has often used this as a basis from which to determine the genre to which the J-corpus may be assigned, while also considering the self-descriptors that are found within it. We will not attempt to ascribe the corpus to a specific category, since modern generic descriptors so often fail to encapsulate medieval conceptions of genre.[33] Rather, we shall examine the self-descriptors which the

[33] On the problems of *post rem* generic classification, see Zumthor (1972), Jauss (1982), and Lacy (1987).

poems contain, which can help us in our search for unity within the corpus, and help us better understand how the texts were seen by their author, as well as their earliest audiences.

From the perspective of the attribution of authorship, what is most meaningful is to ask whether there is any consistency across how the texts describe themselves, and how they are described by their paratexts. If the same descriptors are used in a similar way across the texts, then this can contribute to the corpus' sense of unity, while suggesting that the scribes that copied them may also have felt this unity. Neither of these features is enough to guarantee shared authorship on its own, yet they can contribute to our broader, more holistic picture of the corpus as a unified whole, assisting in the process of attribution.

Morawski (1939: 345) and Munk Olsen (1963a: 12, 1978: xviii–xix) observe three principal terms in use as self-descriptors within the texts attributed to Jehan: *miracle*, *exemple* (also *exemplaire*), and *dit*. Methods of digital textual analysis offer an advantage over those on which previous scholarship had to rely, since they allow us to analyse the forms and distribution of these self-descriptors in a reproducible way. We shall therefore consider each of these terms here, comparing their usage with what we might expect based on an analysis of our broader reference corpora. Following that, we turn to the use of generic descriptors in our more cohesive reference corpora, as well as in the paratexts of fr24432.

3.4.1 *miracle*

The self-descriptor *miracle* appears on a number of occasions in the J-corpus. As a term, it is often used both today and at the time of the texts' composition to describe a body of literature, frequently but by no means exclusively associated with the Blessed Virgin, which makes a moral point using a narrative depiction of events involving divine intervention. Many of the texts of the J-corpus, for example *Respon*, share a great number of similarities with *miracles* such as those of Gautier de Coincy. Indeed, one of the latter's *Miracles de Nostre Dame*

(*Anel*) is found in fr24432 on ff. 315va–316^{vb}, immediately following Jehan's *Beguine*, perhaps suggesting an acknowledgement of similarity between these texts on the part of the manuscript's compiler.[34]

The similarities between Jehan's *Respon* and Gautier de Coincy's *Anel* are stark. In the former poem, a girl is sold to a canon by her father, but when he attempts to bed her, the mention of the girl's name, Marie, is enough to pull the canon out of sin. In return for turning away from sin, the canon is then saved from a hunting accident by Mary. Gautier de Coincy's *miracle* contains a similar story regarding a last-minute turning away from sin, in which a young *clerc* who has promised himself to Mary is reminded of his promise by the presence of the Blessed Virgin in his marital bed, which convinces him to change his ways. Despite differing metres, the epilogues of both texts are striking in their similarity:[35]

Si prions a la Vierge, qui est et fille et mere,
Qu'elle prist son dous Filz — comme tres digne mere —
Que il nous gart d'enfer, celle prison amere,
Et nous otroit la joie qui est bellë et clere.
(*Respon*, vv. 189–192)

Or prion Dieu le fil Marie,
Qu'il nous doint a touz bone vie
Et aprés vie pardurable,
Car celui est bon heritage.
(*Anel*, f. 316^{vb}, ll. 3–6)

Certainly, then, some of Jehan's texts share some characteristics with texts which label themselves as *miracles*, as texts with a narrative focus but a didactic and parenetic function that tell of an act of salvation. And yet, this is not straightforwardly true of all of the texts in the J-corpus, including *Beguine*, the text that immediately precedes Gautier de Coincy's *miracle* in fr24432, which falls into Kleist's (1973: 66–67) category of 'negativen Mirakeln', that is, stories containing miraculous content but in which a sinner is damned rather than saved. Within *Beguine*, the narrator tells of a beguine who falls into league with the Devil and places

[34] For more on questions of compilation, see the codicological appendix on p. 318.
[35] The surviving copies of Gautier de Coincy's text vary in their formulation of the epilogue, and the version found in fr24432 is not noted by Koenig in any of the 10 manuscripts consulted for his edition (1955–1970, ii: 197–204). I have not found a comparable formulation elsewhere.

the Host into a box with a toad, prompting a Eucharistic miracle in which the Host cries out like a new-born child. The beguine's sin is discovered, she is burnt at the stake, and the Devil takes her soul. The text is therefore distinct from the archetypal *miracle* in that it does not have a happy ending for the sinner, but is nevertheless miraculous in its content.[36]

Texts such as *Beguine* certainly have the same didactic potency as the more positive *miracle*-like texts in the J-corpus, and are perhaps most effective in combination with them, offering a clear depiction of the binary choice which sinners apparently face between following the path of righteousness or continuing to follow the path of the Devil. It is noteworthy that *Chevalier et escuier*, which contains the instance of Jehan's self-attribution and which we have so far seen as an archetype of the corpus, offers perhaps the most explicit version of this dichotomy, in that the knight, who refuses to renounce Mary, is rewarded with a miracle, whereas the squire, who chooses to follow the Devil, dies a humiliating death and is damned.

All this is not to attempt to answer the question of whether or not we should refer to the texts of the J-corpus as *miracles*. Genre is of course a thorny question in the Middle Ages, and as we shall see, there are plenty of other generic descriptors which could just as well be used to describe the texts. Rather, it is simply to display the narrative and functional similarities between Jehan's texts and archetypal *miracles* which suggest that it would not be unreasonable for the author to have solely categorised many of his texts as *miracles*, had he chosen to do so.

From the perspective of authorial attribution, a digital analysis of the texts of the whole corpus allows us to gain a global understanding of Jehan's interaction with the term *miracle* in relation to his texts, a task which previous scholarship has only achieved in part. Looking across the whole corpus of Jehan's texts, we find 18 uses of *miracle*, all of which refer to a

[36] Tales resembling this one, we may note, are common also in *exempla*, as discussed on p. 164. Cf., for example, the Eucharistic tales included in the *Index Exemplorum* (Tubach 1969).

greater or lesser extent to an individual miracle within the narrative, but 3 of which can also be said to function as generic descriptors which refer to the text as a whole. These are found in *Narbonne* and *Pecherresse* (twice), and are displayed in Table 3.17.

Table 3.17 – uses of miracle *as a generic descriptor within the J-corpus.*

Text	Line	LH	KW	RH
Narbonne	221	Ceulz et celles qui ont ce	miracle	escouté
Pecherresse	9	Un de ses **biaus**	miracles	vous voudrai raconter
Pecherresse	145	Ce **biau**	miracle	cy a tous vous senefie

Given the fact that, as we have seen, the poems of the J-corpus share features with texts which describe themselves as *miracles*, it is therefore surprising that Jehan makes very little use of the term himself, and only across two texts. The fact that *miracle* appears much more frequently as a term to refer to events occurring within the narrative suggests that Jehan took the term as solely referring to content, rather than as one which had any formal implications. This is reinforced by the fact that Jehan often categorises *miracle*, be it a generic descriptor or a reference to elements of the narrative, as *bel*, with this term characterising the moral, rather than aesthetic, beauty of the narrative itself. In addition the two examples seen above, there are a further four such cases where a form of *bel* appears within two tokens either side of *miracle*, shown in Table 3.18, such that a third of the uses of *miracle* found anywhere in the corpus are accompanied by this word.

Table 3.18 – other uses of miracle *alongside* bel *within the J-corpus.*

Text	Line	LH	KW	RH
Romme	185	La fist moult **biau**	miracle	Nostre Dame Marie
Narbonne	80	Or oez **biau**	miracle	qui oïr le vouldra
Florence	803	Sailli hors sanz main metre le	miracle	fu **bel**
Buef	644	Vous porrez ja oïr un	miracle	moult **bel**

Of course, this use of *bel* is hardly the most innovative construction. We note in passing that *bel* appears to be preferred at the rhyme, as in *Florence* and *Buef*, while *biau* is consistently used

line-internally and appears like an epithet, perhaps included for reasons of metre. That said, its use does appear to offer a unity to certain texts of the J-corpus. In order to reach more assured conclusions on the extent to which this is an unusual feature in Old French texts in general, it is therefore worthwhile considering whether it also appears in our more general reference corpora, namely the P-corpus of the texts in fr24432 not attributed to Jehan, and the large B-corpus which contains texts from a wide variety of generic, diachronic, and diatopic contexts. In so doing, we find that the 33.33% rate of appearance of *bel* in the immediate vicinity of miracle found in the J-corpus is unusually high, since in the P-corpus *bel* does not appear anywhere alongside the 21 instances of *miracle* ($p = 0.00593$) while, even though such examples do appear in the B-corpus, this figure is still only 10.24% (81 of 791 uses of *miracle*, $p = 0.00434$).

Of similar note is the fact that *miracle* in the J-corpus is often associated with verbs of speech and hearing, with 7 instances (38.89%) in which *miracle* is accompanied by one or more instance of the verbs *oïr*, *escouter*, *entendre*, or *raconter*, as shown in Table 3.19. This once again contrasts with the P-corpus, which contains only one such example (within *Anel*), and thus has a rate of such use at 4.76% (1 of 21 instances, $p = 0.00953$), whereas that of the B-corpus is only slightly higher at 8.34% (66 of 791, $p = 0.00022$).

Table 3.19 – uses of miracle *alongside verbs of speaking and hearing within the J-corpus.*

Text	Line	LH	KW	RH
Chanoinnes	402	Or **oez** les	miracles	que Dieus voult demoustrer
Narbonne	80	Or **oez** biau	miracle	qui **oïr** le vouldra
Narbonne	221	Ceulz et celles qui ont ce	miracle	**escouté**
Mescreant	84	Or **entendez**	miracles	se croire les volez
Pecherresse	9	Un de ses biaus	miracles	vous voudrai **raconter**
Buef	644	Vous porrez ja **oïr** un	miracle	moult bel
Abesse	117	Or **oez** les	miracles	que Diex i envoia

Through comparison with these two reference corpora, Jehan's texts therefore achieve a sense of unity due to the fact that the poet associates *miracle* with certain words and actions

more than other authors, but importantly also in the fact that he does not appear particularly interested in calling his texts *miracles*.

3.4.2 *exemple*

Turning to the use of the terms *exemple* and *exemplaire*, we find that, despite a number of uses of the descriptors within the texts, previous scholarship is reasonably dismissive of the possibly that the texts of the J-corpus may be seen to qualify for description as *exempla*. Such texts, which saw a huge rise in popularity following the Fourth Lateran Council in 1215, are short didactic tales which, by the fourteenth century, were frequently told during sermons to illustrate a particular moral argument or to highlight a particular element of doctrine or liturgy (Welter 1927: 1).[37] Such tales would therefore appear to share at least some similarities with the J-corpus, and yet Munk Olsen (1963a: 10) argues that Jehan's texts cannot be described as *exempla*, noting that:

> Les *Dits* de notre recueil ne sont pas sans ressembler aux *exempla* […]. La différence la plus importante, c'est que l'élément narratif, fort réduit dans les sermons pour ne pas dépasser le cadre didactique, s'est amplifié et domine dans les *Dits* ; mais ils servent tous à démontrer une thèse d'ordre moral ou religieux présentée dans leur introduction.

The primary difference between Jehan's *dits* and *exempla*, as Munk Olsen sees it, is therefore primarily one of focus, rather than content. That said, while the surviving collections of *exempla* are indeed reduced in their narrative focus, as for example in the short tales of the *Ci nous dit*,[38] this does not mean that the spoken sermons in which such *exempla* may have been used did not contain the same amount of narrative content as Jehan's poems. The written *exempla* which survive, both in Latin and in the vernacular, were little more than *aide-mémoires* of a broader tale which could then be written up into full stories within a sermon. If, as Munk Olsen himself suggests,[39] Jehan also in some cases developed his tales from sources such as the

[37] See also Brémond *et al.* (1982).
[38] For an edition, see Blangez (1979–1986).
[39] See, for example, Munk Olsen (1978: cxviii).

Ci nous dit, then his work could well be much closer to that of sermon writers than Munk Olsen acknowledges.

Kleist (1973: 67), however, draws a further distinction between Jehan's texts and *exempla*, arguing that there are elements of difference between the two which extend even into the texts' function. The *exemplum*, Kleist argues, 'dient als Illustration einer These, deren Wahrheit in der Predigt schon durch Gründe und Argumente bewiesen worden ist.' For Kleist, *exempla* exist within a much broader framework involving surrounding didactic co-text. Unlike in a sermon, where the *exemplum* comes only after a lengthy exposition of a particular point of morality, liturgy, or doctrine, Jehan's texts have no such accompanying co-text.[40] Indeed, his moral messages appear primarily in the prologue and epilogue to a text, and even then are stated relatively perfunctorily, as a truth universally acknowledged.

Pain, one of the poems within the J-corpus, provides an excellent example of this, beginning with a twelve-line prologue of which the first line is simply: 'Moult fu nez de bone heure qui bien fait en sa vie.' The narrative body of the text makes no explicit reference to this moral message, and yet the story, of a miser who changes his ways, acts wholly in support of the promythium. The message is repeated explicitly only in the eight-line epilogue to the text, of which the first stanza, before a final prayer to Jesus and Mary, reads:

> Qui en sa plaine vie en bien faire labeure,
> Ne se doit pas douter que Diex ne le sequeure ;
> Ains doit douter enfer qui est plus noir que meure.
> Nulz ne scet de sa fin ne le droit jor ne l'eure. (*Pain*, vv. 161–164)

The author, it appears, is explicit in his moral teaching in the prologue and epilogue, and yet is more interested in the main body of the text in implicitly developing his moral argument, rather than explicitly expounding that argument in a way that may be done with an *exemplum*.

[40] On the structure of the medieval sermon, see Wenzel (2016).

Considering these arguments, the overwhelming opinion of previous scholarship appears to be that, while Jehan's poems certainly exist in close proximity to *exempla*, it would be a mistake to describe them as such. For this reason, it is all the more surprising to observe the significant frequency with which the terms *exemple* and *exemplaire* appear as generic self-descriptors in the J-corpus. Across Jehan's texts, there are a total of 37 appearances in various spellings of either *exemple* or *exemplaire*, of which 19 are generic descriptors, as in Table 3.20, with the others either referring to moral lessons given to characters within the narrative, or referring to a specific element of that narrative.

Table 3.20 – uses of exemple *and* exemplaire *as generic descriptors within the J-corpus.*

Text	Line	LH	KW	RH
Respon	186	**Cest**	essemple	nous **moustre** a tous generaument
Chanoinnes	249	Or **entendez** pour Dieu un moult **bel**	essemplaire	
Pommes	5	Je vous weil une	essemple	**amentevoir** et **dire**
Pommes	6	De **bon**	essemple	**dire** ne puet nus estre pire
Romme	5	Un moult **bel**	essemplaire	vous vodrai **recorder**
Deux chevaliers	5	Par **bon**	example	**oïr** se puet l'en **amender**
Enfant rosti	6	Un moult tres **bel**	example	vous vouldrai **recorder**
Narbonne	9	Un moult tres **bel**	example	issy vous en **devis**
Enfant sauva mere	207	En **cest**	example	cy nous est raison moustree
Eaue et vergier	2	Deus	examples	moult **biaus** ou il n'a que reprendre
Eaue et vergier	5	La premiere	example	vus **dirai** sans mentir
Eaue et vergier	97	Le secont des	examples	vous veil **ramentevoir**
Eaue et vergier	104	Par un **moult bel**	example	si con vous **orrez** ja
Mescreant	17	Bone gent je vous veil **moustrer** une	exemplaire	
Merlin	4	Vous **dirai** un	example	a tous generaument
Merlin	217	En **cest**	example	ci doit bien chascun **entendre**
Anelés	4	Un merveilleus	example	ou bon feroit penser
Buef	30	Par un **moult bel**	essample	qu'orendroit vous **dirai**
Buef	790	En **cest**	example	ci vous devez remirer

The use of *exemple* as a self-descriptor therefore appears more consistent across the J-corpus than for *miracle*, with appearances in 13 of the corpus' 24 texts. On a surface level, this would contradict the claims of previous scholarship which, as we have seen, dismiss the possibility that Jehan's texts could ever have been considered *exempla*. However, a more nuanced perspective can be found in the fact that at no point in his texts does Jehan use the Latin word *exemplum*, only French cognates. Indeed, this term does not appear in any of the texts of the

much larger B-corpus, and so it appears a general truth that Jehan's conception of *exemple* and *exemplaire*, and that of other vernacular writers, is not as a direct translation of *exemplum*, but rather a more general term which can be applied to a broader range of narrative texts with a moralising and didactic focus. It is this that can explain the prevalence of *exemple* and *exemplaire* in the J-corpus, despite the generic distance that appears between Jehan's poems and *exempla*.

Turning to the extent to which Jehan's use of *exemple* and its related forms may be said to support a collective attribution of the corpus to a single author, we find that, looking across all 37 appearances of the word, there are a number of collocates that stand out. The phrase *moult tres bel* appears immediately before *exemple* in two cases, forming a construction that does not appear elsewhere in the P- or B-corpora. Meanwhile, the shorter construction *moult bel* appears a total of 7 times immediately before *exemple* in the J-corpus with a rate of 18.92%, while again not appearing in any of the 78 uses of *exemple* in the P-corpus.[41] In the B-corpus, meanwhile, there is only a single instance of *moult bel* before *exemple* out of the 723 uses of the word (0.13%). Looking at the use of *bel* within any structure containing *exemple*, we observe a rate of 31.25% for the J-corpus (10 of 32 instances), 7.69% for the P-corpus (6 of 78), and 1.01% for the B-corpus (8 of 723). There is thus clear statistical significance for the difference in the use of *bel* alongside *exemple* in the J-corpus compared to both the P-corpus ($p = 0.00185$) and the B-corpus ($p = 0$), such that it appears reasonable to indicate that *(un moult tres) bel exemple* was a phrase that Jehan particularly favoured, and which reinforces the claim that the texts containing it are the work of a single author.

The adjective *bon* also occurs 6 times in the J-corpus in the position immediately before *exemple* (16.22%). Comparing this to the P- and B-corpora, we find a comparable rate for the P-corpus (17.95%, 14 instances of 78 occurrences of *exemple*, $p = 0.83887$) but a much lower

[41] There are, however, two instances of *tres bel* appearing before *examplaire* in *Arbre*.

rate of 4.56% for the B-corpus (36 instances of 789 occurrences of *exemple*, p = 0.00456). This is notable in that it suggests that Jehan's use of this phrase was not particularly unusual amongst the texts which were chosen for inclusion in fr24432 (the J- and P-corpora), but that the use of the phrase across the codex as a whole is unusually high compared to the rate found in a broader selection of texts, such as the B-corpus, which contains texts spanning a greater range of periods and dialects.

As for the verbs of speaking and hearing which we found to appear frequently either side of *miracle*, in the J-corpus we observe 16 instances of the use of one or more of *oïr*, *raconter*, *escouter*, *dire*, *recorder*, and *(r)amentevoir* within the immediate context of *exemple*, resulting in a use rate of 43.24%.[42] This time, much lower rates are observed both for the P-corpus (17.95%, 14 of 78, p = 0.00389), and the B-corpus (6.46%, 51 of 789, p = 0), reinforcing the coherence of the J-corpus, and Jehan's apparent interest in presenting the oral nature of his storytelling.[43]

The didactic associations of *exemple* are also made clear by its frequent collocation with words which also have such connotations, including *montrer* (7 occurrences, 18.91%), *prendre* (6, 16.22%), *amender* (3, 8.12%), and *donner* (2, 5.41%). This didactic association of the term is reinforced within fr24432 as a whole through the P-corpus, where we find 9 instances of *montrer* alongside *exemple* (11.54%, p = 0.28559), 11 of *prendre* (14.10%, p = 0.76007), none of *amender* (p = 0.01990), and 4 of *donner* (5.13%, p = 0.95884). That only one such use (*amender*) shows a statistically significant difference from that in the J-corpus therefore suggests that these associations are not unique to the J-corpus, but it is nevertheless a useful observation in helping us to better understand the ways in which *exemple* and *exemplaire* were understood within Jehan's immediate literary context, as terms inextricably linked with didacticism.

[42] The one instance of *amentevoir* in *Pommes* is a hapax in Jehan's corpus.
[43] For more on this, see the section on performance on p. 176.

A similar story appears in the B-corpus, where *montrer* appears alongside *exemple* at a rate of only 4.56% (36 of 789), *prendre* at 12.17% (96), *amender* at only 0.25% (2), and *donner* at 5.70% (45). Again, the rates for *prendre* and *donner* are not statistically distinct from those in the J-corpus ($p = 0.47529$ and $p = 0.93402$ respectively), suggesting a general acknowledgement of the didactic associations of *exemple* and *exemplaire* amongst authors writing in medieval French in a number of locales and periods. It is, however, notable that there are clear, statistically significant differences in the rates for *montrer* ($p = 0.00103$) and *amender* ($p = 0.00027$), suggesting at least some distinction in this area between the texts of the J- and P-corpora and those written by others in different *milieux*.

A final collocate of relevance to *exemple* is *cest*, which appears in the J-corpus directly before *exemple* on 4 occasions, with a rate of 10.81%. In the P-corpus this rate is much lower (at 1.28%, 1 of 78 instances), with a p-value of 0.02581 suggesting a statistically significant difference between Jehan's use of this term alongside *exemple*, compared to those in the remainder of fr24432. In the B-corpus this rate is 4.69% (37 of 789), which gives a p-value which is too high to confirm statistical significance ($p = 0.07479$), but which is nevertheless close enough to our threshold ($p < 0.05$) to support the general trend observed through comparison with the P-corpus, according to which Jehan uses *cest* more frequently alongside *exemple* than might usually be expected.

In this case, the prevalence of *cest* comes perhaps as a result of the fact that Jehan's references to *exemple* frequently appear in the epilogues of his texts, where the poet often chooses to present a strong didactic point about the moral of the whole tale, and thus the use of the deictic *cest* in this position is logical. At the very least, we may say with some certainty that a statistical approach to *exemple* has given a number of clear indications of the ways in which Jehan's texts differ from more diverse reference corpora, supporting our suggestion that these texts may be asserted to be the work of the same author.

3.4.3 *dit*

Thus far, in our examination of the generic descriptors used to describe Jehan's texts, we have omitted the one which appears most frequently in modern scholarship to describe them: *dit*. The term is a fraught one, which does not have a concrete and exclusive definition,[44] and yet it is the term which has come to be used consistently when referring to Jehan's texts. Both Morawski and Munk Olsen use the term to describe Jehan's texts throughout their analyses, with Munk Olsen even going so far as to include the term in the title of his edition of Jehan's texts. It is therefore surprising to find that the generic self-descriptor *dit* is actually used relatively sparsely in the J-corpus, appearing on only 16 occasions (Table 3.21), which makes it slightly less frequent than *exemple*.

Table 3.21 – uses of dit *as a generic descriptor within the J-corpus.*

Text	Line	LH	KW	RH
Respon	4	Weil **recorder** un	dit	dont la matiere est **belle**
Pommes	333	Par essemple vous ai en ce	dit	**raconté**
Romme	2	Veil **dire** un **biau**	dit	qui est sanz vilenie
Chevalier et escuier	3	Que quant je **dis** un	dit	a tous generaument
Narbonne	4	Or vous **dirai** bon	dit	qui ne me fera **taire**
Cordouanier	8	A tiex gens n'ai je cure de **biaus**	diz	**recorder**
Enfant sauva mere	27	**Entendés** un **biau**	dit	que je vous veil **noncier**
Anelés	333	Ainssi com vous **orrés** se tout mon	dit	**recort**
Buef	3	Veul **recorder** un	dit	a touz communement
Buef	33	Et pour ce que mon	dit[45]	soit plus ferme et estable
Buef	36	Qu'a vos ames puist estre ce **biau**	dit	proufitable
Buef	381	Or commence **biau**	dit	et plain de grant sustance
Buef	789	Vous qui avez **oÿ** ce **biau**	dit	**recorder**
Abesse	1	Ceus qui voudront oïr un **biau**	dit	**recorder**
Abesse	6	Un	dit	qu'ai en propos de cest miracle faire[46]
Sauveur	2	Je vous veul **recorder** ung	dit	de grant pitié

[44] The most comprehensive attempt at a description of the *dit* before its later re-use at the hands of Guillaume de Machaut and others is that by Monique Léonard (1996). While her work is impressive in its compilation of a comprehensive list of texts which are considered *dits*, Léonard's description of what constitutes a *dit* extends only to that of a 'pièce rimée, non chantée, plutôt brève, souvent détentrice d'un enseignement ou d'un message à transmettre au lecteur' (1996: 350), which could comfortably apply to a great many other generic categories. For a broader examination of the *dit* from the thirteenth to the fifteenth centuries, see the volume edited by Delage-Béland & Salamon (2022), and in particular the chapter by Gabriel Cholette.

[45] This use of 'dit' could solely be intended to refer to the narrator's discourse, but given its position in the prologue of *Buef*, as well as its proximity to a much less ambiguous example three lines later, its inclusion appears reasonable on the basis that there is at least some generic reference being made here.

[46] Note here Jehan's division between *dit* as a generic marker, and *miracle* as a description of an event, as noted also by Cerquiglini-Toulet (2022: 129, n. 1).

We should note that in the case of *dit*, the extraction of reliable data has been difficult, since RNNTagger in most cases erroneously marks the noun *dit* as a form of the verb *dire*, meaning that for the J-corpus it has been necessary to manually determine which uses of the word form *dit* can be considered generic descriptors, and which cannot.[47] For this reason, given the scale of the P- and B-corpora, which contain 765 and 16,161 instances of the word forms *dit*, *diz*, and *dits* respectively, comparison with all observed forms in both corpora has in this case not been possible. Instead, here I consider the 16 nominal uses identified for the J-corpus in Table 3.21 alongside the 22 instances which RNNTagger has correctly registered as the noun *dit* in the P-corpus, as well as the 510 instances in the B-corpus. While we must be cautious about drawing any firm conclusions from these data, we nevertheless observe similar patterns to those we noted for *miracle* and *exemple*, namely that there is a distinction between the frequency of Jehan's use of particular collocated words, and the frequency with which we would expect to see these across a broader corpus.

Once again, we observe a particularly high frequency of *bel* immediately preceding *dit* in the J-corpus (7 instances of 16, 43.75%), which contrasts with a low rate within the P-corpus (2 of 22, 9.10%, p = 0.01553), and in the B-corpus (7 of 510, 1.37%, p = 0). More strikingly, in 12 of the 16 observations of nominal *dit* in the J-corpus we find at least one verb of speaking and hearing (either *recorder*, *raconter*, *dire*, *taire*, *entendre*, *noncier*, or *oïr*), giving a rate of 75%, compared to the relatively low rate of 13.64% for the instances in the P-corpus in which *dit* is tagged as a noun (3 of 22, p = 0.00010), and 15.10% in the B-corpus (77 of 510, p = 0).

Thus, while Jehan does not use *dit* as often as we might anticipate based on the frequency of the term by modern scholarship to describe his texts, instead making greater use of *exemple*

[47] For reference, only 5 of the 16 manually identified uses of *dit* as a noun in Jehan's texts have been correctly identified by RNNTagger, so we may hypothesise, assuming the rates of failure are the same, that there are some 70 such examples to be found across the P-corpus, and perhaps some 1,632 examples in the B-corpus.

and its variants, the same trends appear in the case of *dit* as for the other generic descriptors we have examined. In particular, *dit* displays high rates of collocation alongside simple, positive adjectives such as *bel*, as well as with verbs of speaking and hearing, like *oïr*. The fact that these similarities apply across the terms used in the texts, and do so in a way that contrasts with the rates observed for our reference corpora, reinforces the collective attribution of the J-corpus to a single author. It also suggests that, for Jehan, the words used to describe his texts were largely interchangeable, particularly *exemple* and *dit*, and that he decided which word to use primarily on the basis of the metrical structure and requirements.

3.4.4 Generic descriptors outside of the J-corpus

Having observed how the uses of generic descriptors within the J-corpus offer a sense of unity to Jehan's texts, it is also worthwhile briefly considering the use of such terms in two of our other reference corpora, namely the L-corpus and the F-corpus, to contrast the approach to generic description found in other single-author and/or single-genre corpora with the J-corpus.[48] Beginning with the L-corpus, comparison is valuable here in that it allows us to observe how another author besides Jehan approached the question of generic self-description within her texts, which are also internally consistent in terms of their genre. Within this corpus, we observe 31 instances of *lai* as a generic descriptor used by Marie de France either to describe her texts, or the sources from which she claims the poems are drawn. The term is particularly frequent in the prologues and epilogues, just as is the case for the J-corpus, but unlike Jehan's texts, Marie's are more consistent in their use of the term, with appearances at least once in every text of the corpus. A further descriptor does appear, with references made to a given text or its contents as a *conte*, but these are much less frequent, and all four instances appear in the immediate vicinity of *lai*, as presented in Table 3.22.

[48] Given the generic and stylistic variation present in Rutebeuf's *œuvre*, the texts of our R-corpus, comparison with this corpus would be of less value here.

Table 3.22 – uses of conte *as a generic descriptor within the L-corpus.*

Text	Lines	LH	KW	RH
Guigemar	19–20	Les	contes	que jo sai verais, / dunt li Bretun unt fait les **lais**,
Guigemar	883–884	De cest	cunte	qu'oï avez / fu Guigemar li **lais** trovez,
Fraisne	1–2	Le **lai** del Fraisne vus dirai / sulunc le	cunte	que jeo sai.
Eliduc	1–3	D'un mult anciën **lai** Bretun / le	cunte	e tute la raisun / vus dirai, […]

In all these examples, Marie positions the term *conte* as subordinate or external to the *lai*, either as an underlying source around which a *lai* was composed (as in the examples from *Guigemar*), or as part of the claimed performance of the poem, with the *lai* an object external to the performance, as in the examples from *Fraise* and *Eliduc*.

The analysis of the generic descriptors used in the L-corpus is therefore pertinent for our examination of the J-corpus in that it displays just how consistent certain poets can be in the generic descriptors which appear in their texts. For Marie and likely also her earliest readers, her texts are consistently *lais*, and other generic descriptors become subordinate to this overarching term. Meanwhile, as we have seen for Jehan, the question of which descriptor to use is a little less clear, with a number of possible terms existing in alternation, and on a roughly equal footing. Comparison with the L-corpus has therefore allowed us to gain a greater sense of the ways in which Jehan's work is distinct from that of other authors.

Turning to the F-corpus, we again find that the use of generic self-descriptors is more consistent than that displayed by Jehan, with the term *fabliau* and its orthographic variants appearing in 11 of the corpus' 22 texts, including 11 uses of *fabliau*, 8 of *fablel*, and 1 of *fable*. *Conte* is used as a generic descriptor for the text at hand on just a single occasion (*Dame Auberee*, v. 2), but its use elsewhere appears to be in opposition to *fabliau*, as indicated by two instances where the two are juxtaposed as examples of the minstrel's craft, with the narrator referring to the 'contes et flabiax' (*Berangier au lonc cul*, v. 1) or 'flabliaus et contes' (*Des trois aveugles*, v. 5) that he knows. Given the fact that both texts also describe themselves as *fabliaux*, these uses of

the term appear to have been included in the text simply to fill out the metre and increase the sense of the performer's knowledge, as well as in the latter case to provide a rhyme. Again, we thus observe that the texts of the F-corpus, despite their varying origins and authors, make more consistent use of a single self-descriptor than the texts of the J-corpus.

This consistency of approach seen across at least half of the texts of the F-corpus therefore reveals much about the relative fluidity of questions of genre in the J-corpus. The authors of the *fabliaux* in the F-corpus are by no means homogeneous in their use of generic descriptors, and yet they are still a good deal more consistent than Jehan. It is therefore Jehan's willingness to use a number of different terms to describe his poems, we discover, that marks him out among other the authors responsible for the texts in our other reference corpora, thus adding an exclusionary unity to the J-corpus, given the diversity it contains.

3.4.5 Generic descriptors in rubrics

One final angle from which the generic descriptors of the J-corpus can be examined is the codicological one. While Jehan appears to have been reasonably fluid in his use of generic descriptors, it is interesting that the compiler(s) and rubricators in fr24432, in which most of his texts survive, were less flexible in their approach. Of the 22 texts from within the J-corpus which are found in fr24432, all but *Pommes* and *Anelés* contain an initial *incipit* marker, and in all 20 cases, the rubricator describes the text as a *dit*. While the same term also appears in the *explicits*, it is less frequent, appearing in 9 of the 21 *explicit* markers (only *Cordouanier* has no *explicit*). Significantly, *Pommes* and *Anelés*, the two texts without the use of the word *dit* in their *incipit*, are given such a descriptor by the *explicit*, meaning that at some point in the paratexts of fr24432, all 22 of the manuscript's texts that we have included in the J-corpus are described as *dits*, and indeed this term is the only generic descriptor used in the paratextual material to

describe any of the texts.[49] This consistency of the use of *dit* is doubtless what explains the tendency of modern scholarship to describe Jehan's texts as such, revealing the power that paratextual information can exert in influencing modern interpretations of texts transmitted in manuscripts.[50] More importantly for our purposes here, it also suggests that Jehan's poems held a certain generic unity in the mind of the compiler(s) of the manuscript, or those of the exemplar(s) from which the texts were copied, as well as at least a certain number of early readers, which may well have resulted from an awareness of shared authorship for these texts, the concrete bibliographical evidence for which has since been lost to us.

The poems were clearly seen by the compiler(s), who may have been among the early readers of the texts, as conforming to a certain text type, and yet we ought to be careful not to read too much into this use of descriptors in fr24432 in particular. The term *dit* is by far the most frequent across the paratexts of the codex as a whole, being used in the *incipits* of a further 32 texts which are not attributed to Jehan,[51] and appearing in the *explicits* of a further 23.[52] Allowing for overlapping usage, we therefore find that, including Jehan's poems, of the 85 texts across fr24432 for which either an *incipit* or *explicit* is included, 56 (65.88%) are at some point described as *dits* in their accompanying paratexts.

The term thus appears to have been in far more frequent use by the compiler(s) of fr24432 than by Jehan himself, and it is therefore clear that the decision made by previous scholarship to ubiquitously describe Jehan's texts as *dits* stems not from the use of the term by the author,

[49] In the paratext surrounding the two texts attributed to Jehan outside of fr24432, *Sauveur* (BnF Arsenal 2115) and *Abesse* (BnF fr. 12483), no such descriptor is used, suggesting this feature may be particular to fr24432.
[50] Given the relative paucity of self-description using the term *dit*, this therefore casts doubt on the validity of the approach taken by Cholette (2022: 93), who selects his corpus of 34 'dits non allégoriques où le caractère narratif domine' (including those by Jehan) by excluding those which are described as something other than *dits* anywhere in the body of the text or in their manuscript paratexts. Selecting texts by this method, based on our analysis many of Jehan's poems should in fact have been excluded from Cholette's corpus.
[51] For ease, here I refer to texts by their item numbers in fr24432, namely items 2, 10, 12–14, 22–23, 43–44, 47–48, 51, 54, 56, 60, 62–64, 66, 68–70, 75–76, 79–83, and 86–88. For titles, see Appendix A: fr24432 contents.
[52] Items 1–2, 5, 12–13, 22, 44, 54, 56, 62–64, 66, 68–70, 79–83, and 87–88.

but from engagement in a chain of reception initiated or followed by the compiler(s) of fr24432. That is not to say that such a decision is invalid, and we do not seek to make a normative claim regarding the correct generic term for describing the texts of the J-corpus, but digital textual analysis has nevertheless revealed that the question of generic description in Jehan's poems is not as straightforward as may be assumed, and depends as much on reception as it does on the preferences of the author.

To conclude this section, we have found across our statistical examination of generic descriptors that Jehan has his own distinct method of describing his texts, which differs from the expectations of previous scholarship, from other authors and genres, and from the perspective of the compiler(s) of fr24432. This distinctiveness, in and of itself, reinforces the shared author hypothesis. Further, the fact that the compiler(s) of fr24432 and subsequent readers have chosen to see Jehan's poems as *dits*, whatever the author himself may have intended and however broad the definition of this term may be, reveals an awareness of the similarities shared between the texts of the J-corpus from Jehan's near-contemporaries onwards, which in itself offers further evidence in support of the claim that these texts share an author.

3.5 Performance

3.5.1 Evidence of orality

We have seen alongside our examination of the use of generic descriptors that Jehan's work is replete with verbs of hearing and speaking, such as *oïr*, *entendre*, *escouter*, *dire*, *raconter*, *recorder*, and *(r)amentevoir*. These were first observed by Morawski (1939), and are not the only hallmarks of orality that we find in Jehan's poems. Further examples often appear in the texts' prologues and epilogues, and include appeals to an audience to come close (e.g. 'Tuit cil soient benei de Dieu et de sa Mere / Qui se trairont en sa et orront ma matere', *Pommes*, vv.

1–2), calls for silence (e.g. '[…] s'il vous plaisoit a taire,' *Mescreant*, v. 18), and the use of terms to directly address an audience (e.g. 'bone gent,' *Chevalier et escuier*, v. 1).

While these features could, as has been established for many other medieval texts, be a mere imitation of orality,[53] their prevalence was enough to convince Morawski, no doubt drawing on the work of Edmond Faral (1910), that Jehan's texts were 'sans doute destinés à la récitation publique, sur une place, un carrefour, le dimanche ou aux jours de foire' (1939: 343). Munk Olsen (1978: xiii) agrees, arguing that the texts appear to have been written to be 'récités en plein air sur la place publique après la messe ou un jour de foire.'

Whether or not we go so far as to take all of these oral features at face value, in the way that previous scholarship has done, it is nevertheless still clear that Jehan's texts are suited to performance. It is therefore not unreasonable to assume that the poems of the J-corpus sit comfortably alongside the vast majority of other short poetic texts from the fourteenth century and earlier, in that they were, at least on occasion, experienced by audiences through an aural medium.

3.5.2 *bone gent*

With this orality in mind, digital textual analysis can give us a greater and more concrete understanding of the interaction of the J-corpus narrator with his audience, particularly in comparison to other texts which are equally likely to have been performed, such as the *fabliaux*. In doing so we gain a clearer picture of the specificity, or rather the specific generality, of Jehan's targeted audience in particular, which goes some way to support Munk Olsen's suggestion (1978: xiii) that the poems were composed to be 'récités en plein air sur la place publique,' and enhances our hypothesis that the texts of the J-corpus share an author.

[53] Cf. Marnette (1998), amongst others.

Beginning firstly with the phrase *bone gent*, which Morawski (1939: 344) identifies as being used by Jehan to refer to his audience, we observe 6 instances of the phrase in this use, as seen in Table 3.23, with a reasonably even spread across texts. Notable here is the frequency with which *bone gent* appears at the beginnings and ends of lines, in an emphatic position that allows Jehan to better engage his audience, but also, particularly in the latter case, one which lends itself to the inclusion of words for reasons of rhyme or metre.

Table 3.23 – uses of bone gent *as a means of referring to the audience within the J-corpus.*

Text	Line	LH	KW	RH
Pommes	217		Bonne gent	or oiez pour saint Pere de Romme
Chevalier et escuier	1	Moult est bone acointance d'estoire	bone gent	
Cordouanier	9	Mais en quelconque leu que je sai	bone gent	
Mescreant	17		Bone gent	je vous veil moustrer une examplaire
Anelés	509		Bonne gent	par un mot qu'elle dist par folie
Anelés	549	"[…] Car du corps ne me chaut"	bonne gent	or oez

Comparing this to the F-corpus of *fabliaux*, the texts of which also contain direct appeals to their audiences, including requests to come close and for silence, it is notable that not a single example of *bone gent* or its variants appears in the corpus. This suggests that this particular term was not necessarily considered a standard means of addressing an audience in a text intended for performance, and that Jehan perhaps included it in his texts as a phrase which he particularly favoured.[54] The texts of the J-corpus are thus somewhat unusual in their use of *bone gent* towards an audience.

There is evidence from within other of our reference corpora, however, that these uses are not unique to the J-corpus. We find 26 instances of *bone gent* across the P-corpus containing those texts found in fr24432 which have not been attributed to Jehan, of which some 12 are used in a vocative sense, as a term of address for an audience. Notably, these uses include one

[54] This is not to say, of course that this phrase is completely absent from the *fabliaux*, only from our select corpus. See, for example v. 33 of *Le Prestre ki abevete* (Noomen & van den Boogard 1983–1998: viii, 299–309).

in *Anel*, the *miracle* by Gautier de Coincy which, as we have seen, shares a number of narrative similarities with the texts of the J-corpus. Most significant, however, is the fact that 12 of these 26 instances are found within the texts of the P1a-corpus, that is those poems written, like the texts of the J-corpus, in monorhyme alexandrine quatrains. The examples of *bone gent* found in these texts therefore constitute some 46.2% of those found across the whole of the P-corpus, despite representing only 11.4% of the total words in that corpus.

It is therefore possible that *bone gent* is a phrase that appears particularly frequently across poems written in the same versificatory form as those of the J-corpus, but there is also another possible explanation worth examining, namely that this is the first indicator in a body of evidence that will grow throughout this chapter that there are texts within the P1a-corpus, in this case particularly *Le dit de Robert le Diable*, which also deserve attribution to Jehan. Of the 12 uses of *bone gent* within the P1a-corpus, 7 of these appear within *Robert*, which on 4 occasions are directed by the narrator towards their audience.[55] The poem tells a reworking of the story of Robert le Diable, the son of a duke of Normandy who commits terrible acts of violence after having been promised to the Devil by his mother, but eventually repents of his sins and is saved. The text therefore aligns closely with the didactic content of the J-corpus, to the extent that we may be led to wonder whether, in the light of this similar use of terminology, it ought also to be attributed to Jehan.[56] While these indicators alone are not enough to reach certainty on the attribution of *Robert* to Jehan, they certainly appear to point

[55] Of the other texts in the P1a-corpus, 4 instances of *bone gent* appear within *Menage*, with 3 of the 4 used to refer to the text's audience. *Menage* is structured as a debate in which three speakers discuss the merits and disadvantages of married life, and as such shares few similarities of content with the J-corpus. It is therefore likely in this case that the multiple uses of *bone gent*, while notable, do not indicate shared authorship.

[56] The only mention of *Robert* in the literature surrounding Jehan de Saint-Quentin is by Morawski (1940: 518), but only as a comparator for a text outside of the J-corpus, with no attention paid to whether the text should be attributed to Jehan. For an edition of the text, see Breul (1895), and for an edition of the earlier *roman* to which it is apparently related, see Gaucher (2006). Should we include it in Jehan's *œuvre*, then *Robert*, at some 1,016 lines, becomes the longest of Jehan's poems. It does nevertheless fit reasonably comfortably alongside *Florence* (832), *Anelés* (780), and *Buef* (796) within the longer grouping of texts from the J-corpus identified on page 151. Indeed, as will be discussed in our codicological examination from p. 318, *Robert*'s positioning within fr24432 (ff. 202ra–215^{rb}) reinforces its inclusion into this 'long-text' grouping of Jehan's texts, given that it appears in the manuscript between *Merlin* (ff. 119ra–201^{vb}) and *Florence* (ff. 215^{rb}–226va).

in that direction, and in our later analysis we will elucidate further areas where the features of *Robert* overlap with the pre-attributed texts of the J-corpus, as a means of supporting a shared attribution.

To return for now, however, to the J-corpus, we find that alongside being unusual in using *bone gent* to refer to their audience, the texts also use this phrase within the narrative itself. By examining the contexts of these 12 such instances, presented in Table 3.24, we can perhaps come to a clearer understanding of whom Jehan may have meant by the term when using it to refer to his audience. The instances of *bone gent* presented here are distinct from the examples used as a term of address for audience members in that they never appear at the very start or end of a line, again supporting the suggestion that the vocative uses of the phrase in the J-corpus are to some extent reflexes used for reasons of metre and/or rhyme. They also differ collectively from the uses directed at an audience in that they appear alongside *toutes* with a frequency of 50%, while no such epithets appear in the data in Table 3.23.

Table 3.24 – cases of bone gent *used other than to address the audience within the J-corpus.*

	Text	Line	LH	KW	RH
A	*Deux chevaliers*	1	A **toutes**	bones gens	ne **devroit** pas desplaire
	Narbonne	1	**Toutes**	bones gens	**doivent** bien oÿr le bien retraire
	Mescreant	1	A **toutes**	bones gens	qui ont entendement
	Pecherresse	10	Que **toutes**	bones gens	**doivent** bien escouter
B	*Enfant rosti*	178	Ainsint sont	bones gens	du vrai Dieu secorus
	Cordouanier	83	De	bone gent	deçoivre est forment convoitous
	Florence	830	Moult de	bonnes gens	sueffrent a grant tort vilonnie
C	*Chevalier hermite*	29	A **toutes**	bones gens	moult d'aumosnes faisoit
	Anelés	292	Devant les	bonnes gens	levez haut le viaire
	Anelés	375	Qu'elle alast en la sale les	bonnes genz	vëoir
	Anelés	751	**Toutes** les	bonnes gens	qui la chose vëoient
	Buef	41	Ou temps qu'il fu jadiz .ii.	bonnes gens	estoient

Jehan's uses of *bone gent* to refer directly to his audience therefore differ collectively from more narrative uses, but they still have a certain utility in helping us to comprehend the apparent possible meanings of this phrase for Jehan. It is possible to distinguish sub-groups within the narrative uses, again represented in Table 3.24. In the first category (A), we find 4 examples

which arguably represent an oblique method of referring to the audience using the third-person, and are thus comparable to the examples in Table 3.23 other than the fact that with these examples, there is a stronger sense of moral obligation attached to the phrases, given the use of forms of *devoir* in 3 out of 4 cases. It is notable also that the uses in category A all appear in prologues, and in 3 cases in the very first hemistich of the text, suggesting again that this was a structure familiar to Jehan, which he may have relied on as a dependable way to begin his texts.

Looking at the second category (B), we find those forms which refer in a general sense to people who are morally 'good,' and yet suffer as a result of sin. From such references, we are reminded of Jehan's claims elsewhere that his audience are 'Pechëeurs pecheresses qui pechent mortelment' (*Buef*, v. 2), thereby setting up *bone gent* as a universalising term for all who sin, and thus all who live.

Meanwhile, in the third category (C), we find the most significant examples for our understanding of what Jehan might have meant when using *bone gent* to refer to his audience. These instances are used to refer to specific characters within the narrative, and can therefore give us a sense of the kinds of people who were deemed worthy of the term. Notably, we observe an apparent contrast between the use of the term in a noble context as a generic phrase indicating members of the aristocracy (the first two instances in *Anelés* and that in *Buef*),[57] and as a term quite clearly referring to poor members of the third estate (*Chevalier hermite*). This, combined with the more general use in the third instance observed for *Anelés*, suggests that the phrase was considered appropriate both for the highest and lowest members of Christian society, as an egalitarian term used regardless of the social status of its referents.

[57] This use appears to differ from that suggested for *bone gent* by Stowell (1908: 155), who claims that the phrase 'is never used by nobles as a title of respect for their fellows of equal rank.'

Jehan's use of *bone gent* to refer to his audience therefore functions in some senses as flattery, in that it is a term that would be appropriate for addressing nobility, and yet on the whole the term is used in a broader sense to apply to his whole audience. The consistency of this use across the texts that contain *bone gent* therefore bind Jehan's corpus together in a way that strengthens the sense that they are the work of a single author. Further, the frequent use of such an egalitarian term also reinforces the universality of Jehan's intended audience, which heightens the possibility that the texts were indeed conceived to be performed, as Morawski and Munk Olsen have suggested, in a public place in front of a crowd that encompassed all subsections of society.[58]

3.5.3 *seigneur*

Alongside Morawski's observation of the use of *bone gent*, Munk Olsen (1978: xiii) also argues that Jehan's texts are characterised by the presence of a number of uses of *seigneur*. Citing Per Nykrog's analysis of the use of *seigneur* in the *fabliaux* (1957: 21), he suggests that the poet's use of the term can tell us something about the intended audience for these poems, namely that it suggests Jehan's audience was a low-status one, with the term used not as a marker of a high-status audience, but rather 'pour désigner un public de vilains ou de bourgeois tout en le flattant un peu' (1978: xiii).

Munk Olsen's interpretation of *seigneur* would therefore certainly contrast with what we have already observed for *bone gent*, in which a term appropriate for nobility is used in a more universalising way, and it does also differ somewhat from that of Nykrog, who suggests that *seigneur* was often used in the *fabliaux* as a general term, referring 'indifféremment à un groupe

[58] The narrative uses of *bone gent* in *Robert* line up with this conclusion for the J-corpus, with one example of type B and two of type C, one of which refers to members of the aristocracy, and the other to pious churchgoers. The poet behind *Robert* appears to have had a similar interpretation of *bone gent* to Jehan, reinforcing the suggestion that they are indeed the same person. A narrative use of the phrase also appears in *Guillaume*, another poem of the P1a-corpus which we attribute here to Jehan, and which we will discuss in the context of *ceulz et celles* from p. 186.

de vilains, de bourgeois ou de nobles' (1957: 21). As we shall see from a digital textual analysis of Jehan's texts, it is Nykrog's interpretation which bears out under scrutiny when applied to the J-corpus, such that Jehan's use of *seigneur* can be said to reinforce the image of his audience gained from *bone gent* as one intended to span all sections of society.

Looking at the data in Table 3.25, we find that Jehan's uses of *seigneur* are notable in their rarity, despite Munk Olsen's insistence on their significance to an understanding of the texts' audience.[59] There are 91 uses of the word across Munk Olsen's corpus of Jehan's *dits*, of which only 3 are direct appeals to the audience.[60]

Table 3.25 – uses of seigneur *as a means of referring to the audience within the J-corpus.*

Text	Line	KW	RH
Pommes	237	Seigneurs	or escoutez que Diex vous soit amis
Juitel	1	Seigneur	je di pour voir qui a entendement
Mescreant	101	Seigneur	cel mescreant dont je vous conte cy

The remaining 88 instances occur within the tales themselves, either during moments of narration (61, 70.45%) or direct discourse (27, 29.55%). Of the examples used in direct discourse, 22 are used as terms of address, which may themselves be useful to our interpretation of what Jehan meant by the term when using it to address his audience.[61] We may separate these examples into two categories of usage: hierarchical and egalitarian. Instances of the former category, found in Table 3.26, offer the use of the term favoured by Munk Olsen, namely as a means of address by an individual of relative social inferiority towards an individual of higher status.

[59] We may choose to repeat one of Munk Olsen's own observations here, made in a critique of Morawski's (1939: 350) conclusions drawn from Jehan's use of religious language, that 'Il ne faut pas oublier que les *Dits* comptent presque 5.000 vers si bien qu'un mot qui se retrouve deux ou trois fois n'est pas nécessairement un mot "que l'auteur affecte beaucoup"' (1963b: 155, n. 17).
[60] It is worthy of note that *Robert* matches this usage, containing 10 uses of *seigneur* of which none appear in a direct address to the audience. *Guillaume* contains a further 20, which again do not appear in a vocative sense.
[61] Of the 5 uses of *seigneur* within direct discourse which are not used as terms of address, there are 2 references to God (*Juitel* and *Mescreant*), 2 to a woman's husband (*Florence* and *Anelés*), and 1 to a speaker's *suzerain* (*Pommes*).

Table 3.26 – hierarchical uses of seigneur *within the direct discourse of the J-corpus.*

Text	Line	LH	KW	RH
Narbonne	117	“	Seignor	” dist li varlez “Or oez que dirai
Narbonne	125	“[…] Encor vous pri	seigneur	que un po m’escoutez
Merlin	55	“A mon	seigneur	Merlin je me met au chemin
Merlin	61	“Mon	seignour	granz mercis je vendrai volentiers
Merlin	91	“Ha mon	seigneur	Merlin ou m’esperance est toute
Merlin	98	“A mon	seigneur	Merlin que je fusse prevost
Merlin	101	“Mon	seigneur	grant mercis ci a noble secours
Merlin	166	“[…] Doucement m’apellaz ‘ha mon	seignour	Merlin’
Anelés	94	“Mon chier	seigneur	par amour vous requier

In the category of egalitarian uses of *seigneur* in Table 3.27, meanwhile, we observe the use of *seigneur* as a polite term which shows respect for one’s interlocutor(s), but not in a way that acknowledges any difference in social status. In some examples, such as the latter two uses in *Respon*, the use of the term is in fact inverted from what Munk Olsen appears to indicate, with *seigneur* used by a suzerain to refer to his vassals.

Table 3.27 – egalitarian uses of seigneur *within the direct discourse of the J-corpus.*

Text	Line	LH	KW	RH
Respon	150	Et dient l’un a l’autre “	Seingneurs	alons nous en
Respon	171	“Amis	Seingneurs	je sui vos sire par Dieu de paradis”
Respon	176	“	Seingneurs	” dist le chanoinne, le roi de maesté
Chanoinnes	33	“	Seigneurs	par celle foi que vous a Dieu devez
Chanoinnes	47	“	Seigneur	” dist li ainsnés “Entendez mon recort
Chevalier et escuier	50	“	Seignor	comment vous est” ce dit li ennemis
Florence	103	Quant son frere leur dist “	Seignours	lessiez ester”
Florence	122	“	Seignours	” dist il “Pour Dieu la vengance prenez
Anelés	221	“Biau	seigneurs	vous veés ci endroit en present
Anelés	273	Et dist “	Seigneurs	je jure par les sains qui sont ci
Anelés	309	“	Seigneurs	or entendez pour Dieu granz et petiz
Anelés	410	“	Seigneurs	ne savés pas pourquoi vos ai mandés.
Anelés	450	Et dist “	Seigneurs	ma fame est de vostre lignage

Comparing the two categories, it is notable firstly that there are a greater number of egalitarian than hierarchical uses of *seigneur* in Jehan’s presentations of direct speech (13 of 22). Indeed, 6 of the 9 hierarchical uses of *seigneur* appear within *Merlin*, a text which revolves almost entirely around questions of respectful means of addressing one’s social superior, telling a story of how a *vilain*’s ingratitude towards his patron Merlin results in his fall from

grace. 'Seigneur' is the first form of address used by the *vilain*, when he is at his most polite, and before the fall in register takes place in his later speech, and so it is unsurprising that the term would be used in this way so often within the text. We can therefore argue that the data are skewed by the inclusion of *Merlin* in the corpus, and thus that Jehan's primary understanding of *seigneur* is as a polite, but fundamentally egalitarian term.

Relevant to the question of Jehan's uses of *seigneur* to address the audience is also the fact that his egalitarian uses within direct discourse are exclusively in the plural, while in the hierarchical uses of *seigneur* the only plural forms are the two appearances of the term in *Narbonne*, and the others are singular. When Jehan uses *seigneur* in the plural to refer to his audience, then, we may argue that he does so for egalitarian purposes, treating his entire audience equally, regardless of their social status. Ultimately, this contrasts with Munk Olsen's interpretation of the term, and once again reinforces the unity of the J-corpus through the suggestion that Jehan, in his texts, is not interested in isolating groups of possible listeners, rather he is keen for any and all individuals to come and hear his moral lessons.

As for how unique this feature is to Jehan's texts, we find that the F-corpus further supports Nykrog's claim, itself made of the *fabliaux*, that *seigneur* can be used 'indifféremment à un groupe de vilains, de bourgeois ou de nobles' (1957: 21), while also supporting our observation that in short narrative texts, *seigneur* is not particularly widespread as a term of address for one's audience.[62] Of the 22 instances that appear across the corpus, only one is used as directly by a narrator, in *Des trois dames qui troverent un vit*, in which the narrator, as part of his parodic epimythium, proclaims that 'seignor, je vos chasti' (v. 116). As with the J-corpus, while the 9 examples appearing within direct speech in the F-corpus to address an interlocutor vary from distinctly hierarchical uses, such as addresses to God, to sardonically

[62] Anecdotally, at least, it does appear to be more common in the *Chansons de geste*.

polite non-hierarchical uses against a group of ne'er-do-wells, the non-hierarchical uses are all in the plural. We thus gain a similar interpretation to that found in the J-corpus, namely that plural uses of *seigneur* are polite, but can be addressed to individuals across society, whereas singular uses may be more hierarchical. Thus, while the use of *seigneur* observed in the J-corpus does little to reveal a unique approach to the term within Jehan's texts, it has nevertheless allowed us to support Nykrog's more generally-applicable claim from a more concrete, falsifiable standpoint, which in turn allows us to gain a more nuanced understanding of the fact that Jehan's texts are united in their attempts to apply to all members of society, regardless of social status.

3.5.4 *tous et toutes* and *ceulz et celles*

This universality of Jehan's intended audience, inclusive of all who wish to listen, is emphasised perhaps most of all within his direct addresses to his audience by his use of the phrases *tous et toutes* and *ceulz et celles*, given in Table 3.28, which reinforce that his moral messages are for the betterment of both sexes, and therefore of society at large.[63]

Table 3.28 – uses of tous et toutes *and* ceulz and celles *as means of addressing the audience in the corpus of texts attributed to Jehan by Munk Olsen (1978).*

Text	Line	LH	KW	RH
Chanoinnes	409	Si prions	touz et toutes	a la Vierge Marie
Narbonne	5	Vous savez	tous et toutes	qui sert sainte Marie
Narbonne	221		Ceulz et celles	qui ont ce miracle escouté
Cordouanier	149	Or prions	tous et toutes	la Dame glorieuse
Juitel	161	Or prions	tous et toutes	de bon cuer le Seigneur
Enfant sauva mere	25	Entre vous	tous et toutes	je vous pri et requier
Eaue et vergier	6	Or metez	tous et toutes	les cuers a retenir
Mescreant	201	Si prions	tous et toutes	a la Virge Marie
Florence	4	Gart trestouz	ceulz et celles	qui nous voudront oïr

[63] Note also the variant 'Touz ceus et toutes celles', used in an oblique reference to Jehan's audience in the first line of *Beguine*, and also in *Buef* (v. 781), as well as the comparable phrase 'chascuns et chascune' in *Chanoinnes* (v. 2), *Povre chevalier* (v. 159), and *Eaue et vergier* (v. 162). Examples of such universalising phrases are thus found in 11 of the 24 poems of the J-corpus.

The position of all but one of these phrases in the final syllables of the first hemistich, often following the phrase *si/or prions*,[64] suggests that this is a formulaic phrase used to fill out a line, but it also aligns nicely with the didactic and parenetic intentions behind Jehan's texts, given that his objective is apparently the edification of all. What is particularly surprising about this construction is that no examples of the same phrase appear in the F-, L-, or R-corpora, and while there are some 38 examples in the B-corpus, in the J-corpus *tous et toutes* and *ceulz et celles* are used exclusively as terms of address for the audience, whereas in the B-corpus just one example is used in this way, from the late-fifteenth-century *Ressource de la chrestienté*. Jehan therefore appears to be particularly unusual in his use of these terms to refer to his audience.

It is all the more surprising, then, that a similar such use should be found in the P-corpus, outside of the texts which have been attributed to Jehan. The phrase *ceulz et celles* appears in the second line of *Guillaume*, the first text in fr24432, where the narrator exclaims 'Dieu gart touz ceulz et celles dont seray entendus' (f. 1ra).[65] This use of the phrase is therefore our first indication that this text could also sit alongside *Robert* as meriting attribution to Jehan. As with *Robert*, we find that *Guillaume* shares a great many similarities with the texts of the J-corpus, both in form and content. Like the texts of the J-corpus, the text is written in monorhyme alexandrine quatrains, and like *Robert* and at least also *Florence* it is a reworking of a longer epic.[66] The text is heavily didactic in its content, telling the story of William, King of England, who undergoes many trials imposed on him by God including exile to France, before eventually proving his piety and being reunited with his family. Certainly, the presence of a single construction is far from enough to guarantee that this text is the work of Jehan, but the

[64] Morawski (1939: 330) also observes as particularly frequent use of this phrase in Jehan's epilogues.
[65] Given that the version of the text to which I refer is my direct encoding from fr24432, for all texts of the P-corpus, including *Guillaume* and *Robert*, I cite folio references rather than line numbers, to avoid conflicting with existing critical editions.
[66] The earlier version has been attributed, with debatable validity, to Chrétien de Troyes. For more on this question of the authorship of the longer epic, see Gowans (1990) and Slautina (2012), who attempts a digital approach to attribution through less directed methods than we use here. For an edition of this original text, see Short (2021), and for an edition of the alexandrine reworking contained in fr24432, see Buzzetti Gallarati (1990).

rarity of this phrase as a means of addressing an audience across all our reference corpora is certainly an indication that shared authorship is possible. As with *Robert*, we will therefore return to *Guillaume* later in our analysis when we are able to shed further light on this question.[67]

Statistical and digital methods have therefore helped us to gain a greater understanding of Jehan's intended audience through analysis of the words he uses, in ways that have both reinforced and caused us to bring into question more qualitative interpretations of the texts. We have, as a result, been able to come to firmer conclusions regarding Jehan's intended audience and mode of reception, namely that Jehan's texts appear most likely to have been intended for oral performance, and were directed at as broad a range of social groups as the performer could convince to listen.

Alongside this, and as in previous sections of our analysis, an important observation for our overarching question of the attribution of authorship remains the unity brought to the J-corpus by the use of certain forms of address. The consistency we have seen throughout in Jehan's approach to his audience ultimately allows us to be more confident in our assertion that the texts of the J-corpus which contain these features likely all share a single author.

Perhaps most excitingly of all, the analysis within this section has also alerted us to two further texts which, as we shall see throughout this chapter, are worthy of inclusion in Jehan's *œuvre*. The evidence given so far is not enough on its own to prove shared authorship, but even by identifying these texts as possible candidates for inclusion, we have also shown the value of the

---

[67] *Guillaume* contains some 948 lines of verse, and is therefore once again longer than the texts of the J-corpus, but not so long as to preclude attribution, being roughly comparable in length to *Florence*, *Anelés*, and *Buef*. *Guillaume* has a 36-line prologue, which as with *Robert* (15) is shorter than that seen for *Anelés* (84) and *Buef* (40), but longer than that of *Florence* (12), while its short epilogue (1 line), equal in length to that of *Robert*, matches our observation above that epilogue length in Jehan's texts shares little to no statistical relation to the length of the rest of the text. Both *Guillaume* and *Robert* are described as *dits* in the manuscript paratexts of fr24432 (*Guillaume* in the explicit, *Robert* in the incipit and explicit), and each contain 2 uses of the generic self-descriptor *dit* (*Guillaume*: ff. 1ra and 9^{rb}; *Robert*: ff. 202ra and 208^{rb}). Neither contains any use of *exemple* or *miracle*.

bespoke statistical and digital approaches presented in this analysis as a means of identifying possible shared authorship, with implications which reach more broadly into other research questions concerning the attribution of medieval texts.

3.6 Versification

From the earliest attribution of texts to Jehan, the primary motivating factor for doing so appears to have been a feeling amongst scholars that the texts could be connected 'wegen der ihnen eignen Einfalt des Vortrags und des unvollkommenen Sprach' (Gröber 1888–1902, ii: 929). In particular, scholars have taken issue with Jehan's apparent 'médiocrité technique' when it comes to versification (Munk Olsen 1978: cxxxvi), using the apparent ubiquity of his failings across the J-corpus as grounds to suggest that the texts share a single author.

Evidence for this comes, so the argument goes, from Jehan's copious repetition of the same word in more than one rhyming position in a stanza (Morawski 1939: 334), as well as in his use of non-elision of schwa in hemistich-final position, particularly with the particles *je*, *se*, *ne*, and *que* (Munk Olsen 1978: cxxxix), and his use of so-called 'easy' rhymes such as words ending in *-ment*. Morawski (1939: 335) also brings our attention to the fact that in two cases (*Deux chevaliers*, vv. 205–210; *Chevalier et escuier*, vv. 69–72b), there are stanzas of more than four lines, the latter of which certainly does not appear to be the result of scribal corruption.

And yet, even if we were to indulge in assertions of aesthetic value, we would still be forced to note, as Morawski does (1939: 335), that '[t]outes ces irrégularités se retrouvent, souvent à un degré plus élevé encore, dans d'autres poèmes de l'époque, écrits dans la même forme métrique.' If, across similar surviving texts from the period, 'la plupart des spécimens […] ne se distinguent guère par une technique souveraine' (Munk Olsen 1963b: 157), then the quality of Jehan's versification is ultimately of minimal value to the question of attribution, since it does nothing to distinguish his work from that of other authors.

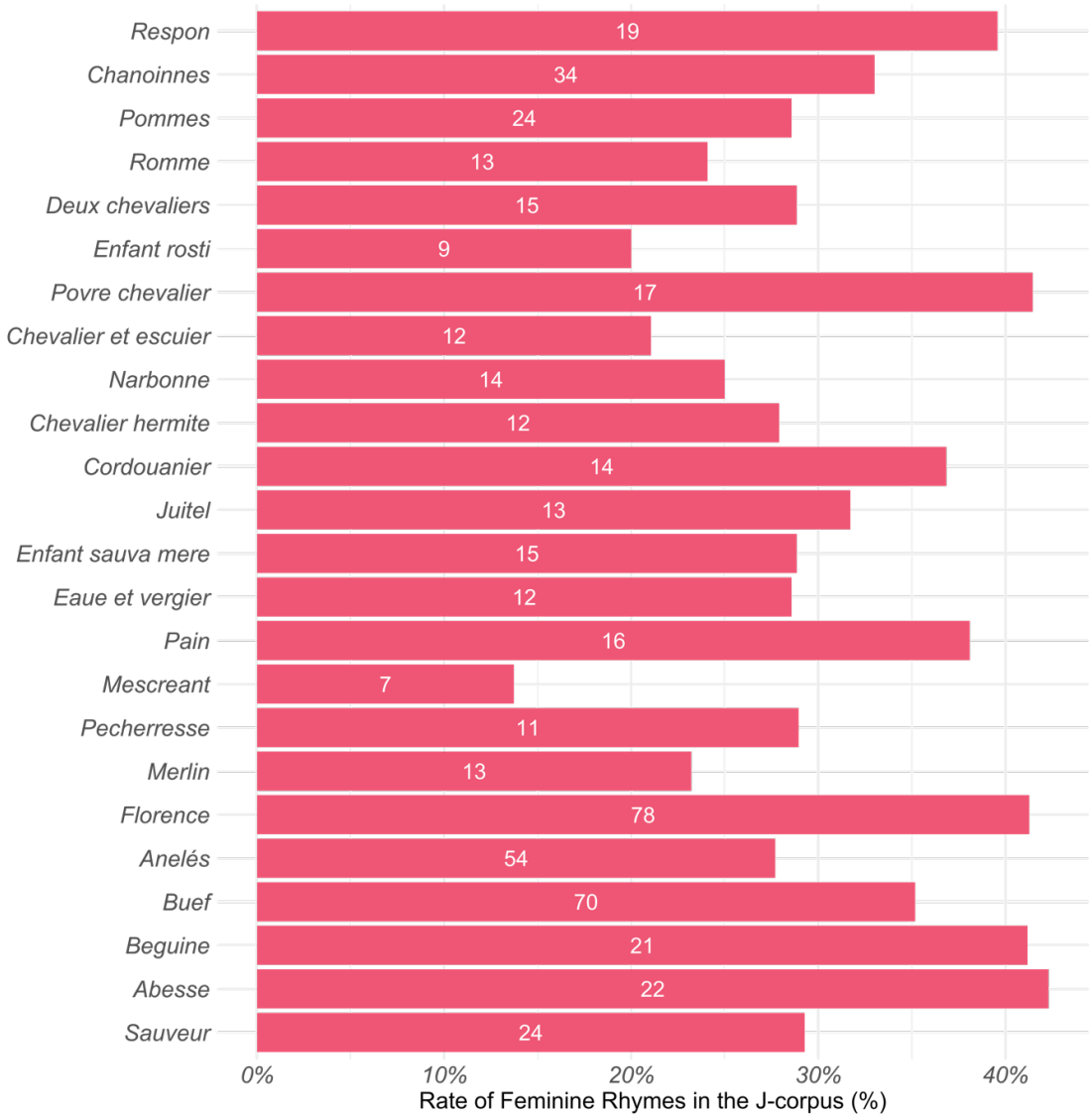
Not only would it be questionably subjective to class Jehan's poetry as either 'good' or 'bad,' it should go without saying that an individual author's output will never be consistent in quality, as they hone and develop their craft across their career, or indeed fall from earlier creative heights. The avenues pursued by previous scholarship regarding Jehan's versification therefore offer little evidence in either direction concerning the collective authorship of the J-corpus. By turning towards digital and statistical analyses of versification, however, we can both more comprehensively interrogate the validity of certain of the existing claims concerning Jehan's versificatory tendencies, as well as pursue new research avenues, which can tell us much more than previous scholarship has been able to determine about the extent to which the texts of the J-corpus form a coherent group, distinct from texts found elsewhere.

3.6.1 Masculine and feminine rhymes

With this in mind, we begin our application of digital methods by considering the use of masculine and feminine rhymes within Jehan's poetry.[68] The mean rate of use of feminine rhymes across the corpus is 31.5%, almost exactly the 1:2 ratio identified by Morawski (1939: 335) for his smaller corpus of Jehan's *dits*, which he claims is 'la norme' for texts of this period. As we shall see, the frequency with which Jehan uses feminine rhymes is not identical to that of poets whose works his texts appear alongside, but it is within the expected range for the fourteenth century, given Lote's (1949–1996, ii: 111–124) observation that the rate of use of feminine rhymes increases from near to zero in the earliest surviving poems in French, to near-parity with masculine rhymes by the end of the fifteenth century.

[68] A feminine rhyme is one ending in atonic *e*. In practice this includes forms such as the third person plural *-ent*, as well as words ending in *-es*. For this analysis, stanzas are said to have a feminine rhyme if they contain at least two forms ending in any of *-e*, *-es*, or *-ent*. To avoid including masculine rhymes ending in *-ent*, stanzas containing a single example of a form ending in *-ment* or *-ant*, or those with a single letter before *-ent* (such as *vent*), or two if the *-ent* is preceded by a vowel (such as *vient*), are excluded on the assumption that *-ent* in these cases is likely to result in a nasal vowel. These requirements are been represented through regular expressions, namely that at least two lines of a stanza must satisfy `es?$|[a-zéäëïöü]{2,}ent$`, and no line in the stanza can satisfy `ment$|ant$|^[a-z]i?ent$|és$`. This will lead to the exclusion of rare stanzas containing third-person plural forms with *m* preceding the *-ent*, such as *entament*, but sampling suggests that the impact on the data is minimal.

Figure 3.52 – rates of use (x-axis) and raw numbers (figures in white) of stanzas containing feminine rhymes for each text of the J-corpus.



However, looking at the rates for individual texts, represented in Figure 3.52, we observe variability in the use of feminine rhymes across the J-corpus. Morawski identifies unusually low rates in *Chevalier et escuier* and *Merlin* (from our analysis, 21.1% and 23.2% respectively), but misses the similarly low rate of 20% in *Enfant rosti*, and the much lower rate displayed in *Mescreant* (13.7%), while also not recognising the much higher than average uses in *Povre chevalier* (41.5%), *Beguine* (41.2%), *Respon* (39.6%), and *Pain* (38.1%). Since he does not include either text in his corpus, Morawski also does not find the highest rate of use to appear within our J-corpus, namely *Abesse* (42.3%), as well as the high rate observed for *Florence* (41.2%).

We may hypothesise that some explanation for this increased use of feminine rhymes may result from the fact that, except for *Pain*, all of these texts displaying high rates of feminine rhymes centre around the experience of women,[69] making words ending in *e caduc* more likely, through for example the increased prevalence of nouns such as *fame* (e.g. *Povre chevalier*, vv. 102, 104), as well as variable adjectives and past participles with feminine agreement (e.g. *Povre chevalier*, vv. 114–115). The inverse argument could be made for the low rate of use found in *Mescreant*, a text containing only male characters which is therefore unlikely to require as many words ending in *e caduc* as some others. This offers a somewhat convincing argument for these texts, but does not appear to be the sole explanation for the variation across the J-corpus, given the absence of female characters in *Pain* (38.1%), which nevertheless has a much higher rate of feminine rhymes than *Romme* (24.1%), *Narbonne* (25%), and *Pecherresse* (28.9%), which centre around a single female character, and yet in turn have lower rates of feminine rhyme use than the corpus average.

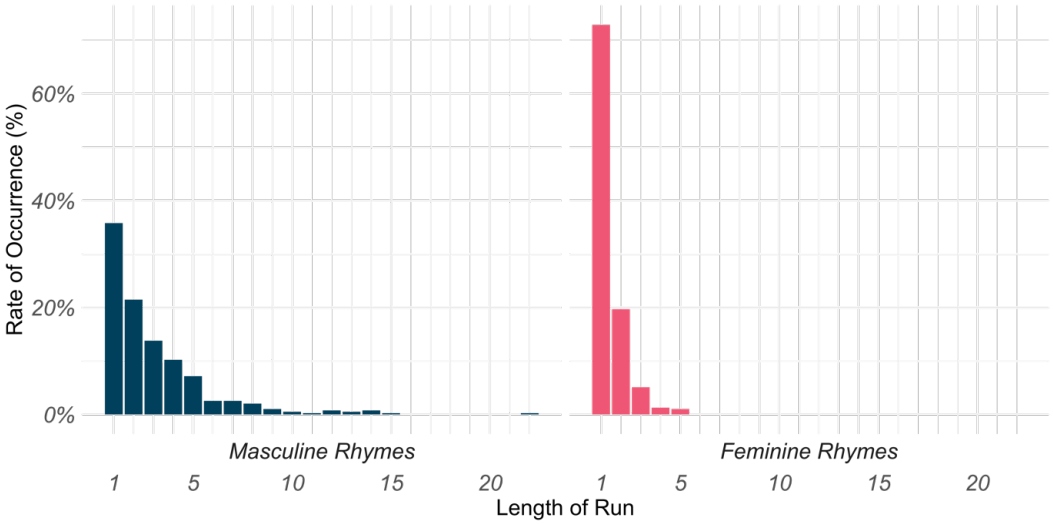
On the whole, then, Jehan's texts fit the overall frequency of use of feminine rhymes observed more widely by Lote and suggested for the corpus by Morawski, but they are certainly more variable than the latter acknowledges. While in some cases this variability can be explained with reference to the contents of a particular text, this is not always the case.

Munk Olsen (1978: cxxxix) argues that within Jehan's texts, monorhyme stanzas containing masculine rhymes appear alongside those containing feminine rhymes 'sans principe fixe.' On the whole, this is true, and there are no texts in which a clear pattern of usage is maintained the whole way through. Nevertheless, there are some tendencies to be observed in the J-corpus. Firstly, on a purely statistical level, by comparing the lengths of sequences, or 'runs,' of stanzas containing masculine versus feminine rhymes, as represented in Figure 3.53, we

[69] In *Povre chevalier*, it is the actions of the knight's wife and of Mary which form the crux of the story.

note that the appearance of stanzas containing rhymes of a given gender flanked on either side by rhymes of the opposite gender, rather than in a run of more than one such stanza, is far more common for feminine rhymes (72.9%) than it is for masculine rhymes (35.8%). This is perhaps at least somewhat as a result of the simple fact that there are fewer stanzas with feminine rhymes than those with masculine rhymes across the corpus. Were this to be the only explanation, however, we might expect runs of 2 stanzas to be the most frequent type for masculine rhymes, with masculine and feminine rhymes alternating on a two-to-one basis.

Figure 3.53 – rates of occurrence for runs of stanzas containing masculine and feminine rhymes across the J-corpus.



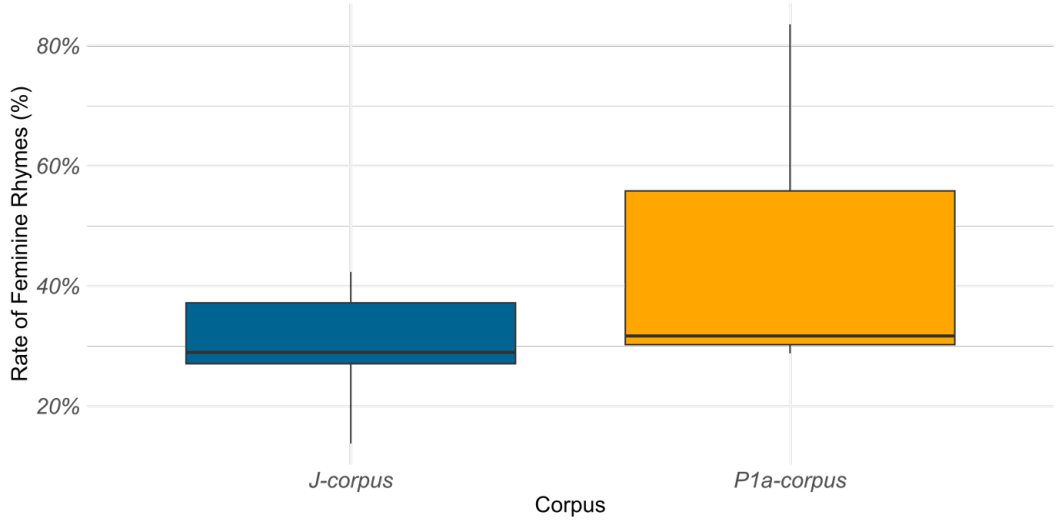
This is not the case for the J-corpus, however, since in fact patterns of one-to-one alternation between stanzas of masculine and feminine rhymes are more common. The greater difference between the use of masculine and feminine rhymes can instead be accounted for by the fact that the longest run of feminine rhymes is 5 (twice in *Florence*, once in *Buef*, and once in *Abesse*), while for masculine rhymes it is 22 (*Romme*): in his texts, Jehan appears far more willing (or perhaps able) to use long runs of stanzas containing masculine rhymes than he is to make use of long runs of stanzas involving feminine rhymes.

Indeed, if we exclude those runs of stanzas containing masculine rhymes which have a length greater than the longest run of feminine rhymes (5), we find that only 770 stanzas of a total of

1,172 fit these criteria, putting the stanzas with masculine rhymes on a far more even footing with stanzas containing feminine rhymes, of which there are 539. For the most part, then, the two-thirds predominance of masculine rhymes across the J-corpus perhaps results from an aesthetic choice with Jehan preferring, if not sticking rigidly to, either a one-to-one or a two-to-one alternation between stanzas containing masculine and feminine rhymes, but permitting the occasional interruption by longer runs of stanzas containing masculine rhymes.

By comparing the rates of use of feminine rhymes in the J-corpus with the P1a-corpus of texts found in fr24432 which have not been attributed to Jehan, but which are also written in monorhyme alexandrine quatrains, we can gain a clearer understanding of whether the 2:1 ratio of masculine to feminine rhymes is as typical as earlier scholarship has suggested.[70]

Figure 3.54 – box and whisker plots for the rate of use of feminine rhymes in the J- and P1a-corpora.



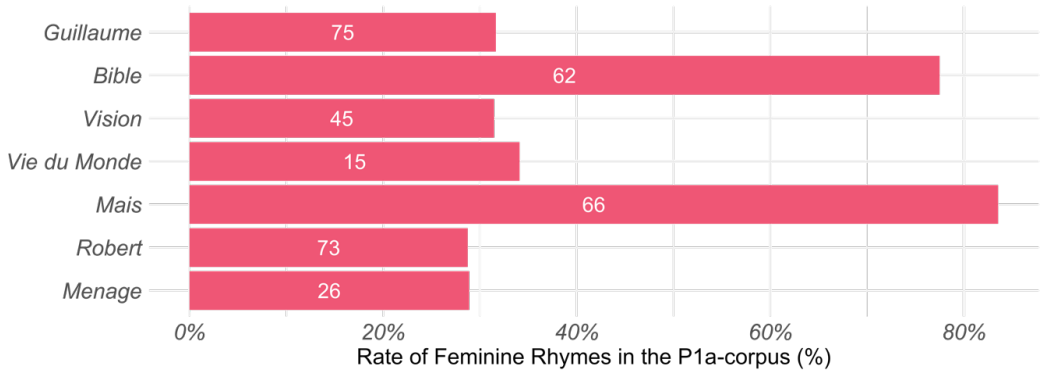
From Figure 3.54, we observe that while the median rate of use of feminine rhymes in the J-corpus (28.9%) is relatively comparable to that of the P1a-corpus (31.6%), the range is vastly

[70] The criteria for determining whether a stanza within a text of the P1a-corpus contains feminine rhymes are the same as were defined for the J-corpus in note 68. Given the non-stanzaic form of many of the texts within the broader P1-corpus, an analysis of this corpus would require us to examine lines rather than stanzas, meaning that the criteria for determining whether a rhyme is feminine are less reliable. We therefore limit our scope for this subsection to the P1a-corpus, although rates of use of feminine rhymes in the P1-corpus will be examined in the next sub-section when this provides a useful point of comparison to other versificatory features.

different, with the lowest rate in the P1a-corpus only slightly lower than the median rate in the J-corpus (28.7%), while the highest rate in the P1a-corpus reaches a rate of 83.5%, almost twice the highest rate of the J-corpus (*Abesse*, 42.3%). The middle 50% of the data in the J-corpus are also spread over a smaller range, with an inter-quartile range of 10.1% in the J-corpus, but 25.6% for the P1a-corpus, despite the P1a-corpus containing fewer texts.

More light can be shed on these observations by looking individually at the texts of the P1a-corpus, as presented in Figure 3.55, where we find that there are two significant outliers within the P1a-corpus in terms of their use of feminine rhymes, with the other texts relatively comparable. The presence of these two outliers, *Mais* (83.5%) and *Bible* (77.5%), reinforces the suggestion that significant variation existed in the use of feminine rhymes in texts circulating in a similar milieu to the texts of the J-corpus, even amongst texts with the same versificatory form, which would arguably result at least partly from authorial preference.

Figure 3.55 – rates of use (x-axis) and raw numbers (figures in white) of stanzas containing feminine rhymes for each text of the P1a-corpus.

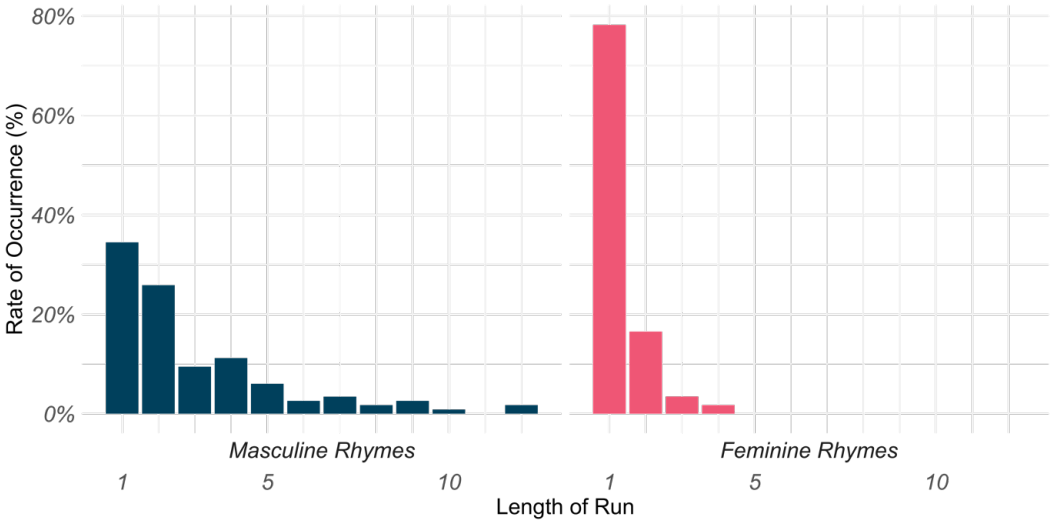


Reinforcing this sense is the fact that neither *Mais* nor *Bible* contain content which accounts for their high rate of feminine rhymes. One may be tempted in the case of *Bible* to argue that the high rate of feminine rhymes is explained by the text's focus on Mary and other biblical women, which might give the poet a greater range of nouns and adjectives containing atonic *e* with which to produce feminine rhymes, and yet while this does occasionally appear to ring true, in a greater number of instances it is not the case. A number of feminine rhymes appear

on words which bear no relation to the women referenced in the text, and thus, as we saw for the texts of the J-corpus, a focus on female characters may sometimes increase the use of such words and therefore increase the rate of feminine rhymes, but this is not always guaranteed.

The rate of use of feminine rhymes, then, appears to be at least in part a feature of authorial preference, with significant variability possible even between texts of the same form. It is therefore all the more notable that *Guillaume* and *Robert*, the two texts for which we have already observed a number of significant similarities with the texts of the J-corpus, should display rates of use of feminine rhymes that place them squarely in the middle of the rates observed for the J-corpus, at 31.6% and 28.7% respectively. As a result, we are able to strengthen our hypothesis that these texts may in fact also be the work of Jehan. The presence of similar rates in other texts of the P1a-corpus, however, namely *Vision*, *Vie du Monde*, and *Menage*, of which the second is at least tentatively attributed to Rutebeuf,[71] means that this evidence on its own is not enough to confirm attribution, and instead must feed into the more holistic evidence which we are building for these texts over the course of this chapter.

Figure 3.56 – rates of occurrence for runs of stanzas containing masculine and feminine rhymes in Guillaume *and* Robert.



[71] For example by Zink (1989–1990, ii: 450-458). Bastin & Faral (1959–1960) and Jubinal (1874–1875), responsible for earlier editions of Rutebeuf's *œuvre*, are more convinced of the text's attribution.

Returning also to the 'runs' of stanzas containing masculine versus feminine rhymes, as represented in Figure 3.56, we observe a similar pattern in *Guillaume* and *Robert* to that seen for the J-corpus, namely that runs of masculine rhymes are often much longer than runs of feminine rhymes, with two 12-stanza runs of masculine rhymes appearing in *Robert* and one of 10 stanzas in *Guillaume*. The longest run of feminine stanzas in either text is a single 4-stanza run found in *Guillaume*. As with the J-corpus, by excluding those runs of stanzas containing masculine rhymes which exceed a length of 5, the number of masculine stanzas in both texts drops from 343 to 220, and is thus much closer to the same figure for the number of feminine stanzas (148), therefore suggesting that for both *Guillaume* and *Robert*, as in the J-corpus, stanzas of masculine and feminine rhymes alternate on a roughly one-to-one or two-to-one basis, with the overall data skewed by the presence of longer runs of masculine stanzas.[72]

To end our section on feminine rhymes, then, we observe that the usage observed for the J-corpus is arguably close to typical, as previous scholarship has suggested, and displays some variability within these rates. Evidence of far more variability within other texts written in monorhyme alexandrine quatrains, however, supports the collective attribution of the J-corpus to a single author, as well as reinforcing the possibility that *Guillaume* and *Robert* are also the work of Jehan. Some level of variation in the use of feminine rhymes therefore appears to be possible both between the work of different authors, and within the *œuvre* of one single author. We have seen that while we are occasionally able to explain this variation by considering a text's content, it is frequently impossible to do so.

[72] Notably the same is not true of the remainder of the P1a-corpus. *Mais* contains a run of 18 feminine stanzas, and even excluding *Mais* and *Bible*, the rate of occurrence for runs of 3 masculine stanzas (21.1%) exceeds that observed for runs of 2 (19.3%), going against the trends observed in the J-corpus and for *Guillaume* and *Robert*.

3.6.2 Rhyme frequency and 'easy' rhymes

Many of the value judgements concerning the quality of Jehan's poetry stem from the claim that Jehan makes excessive use of so-called 'rimes faciles' (Munk Olsen 1978: cxxxvii), the preponderance of which allows us, so the argument goes, if not to identify the work of an individual author, then at least to 'distinguer le bon poète et le versificateur irrémédiablement mauvais' (1963b: 158). Morawski (1939: 334) defines this class of 'rimes faciles' as consisting primarily of nouns and adverbs ending in *-ment*, as well as verb forms sharing common endings, in particular the present participle (*-ant*) and infinitives.

While we may wish to dismiss the value judgements with which they are associated, the supposed abundance of these so-called 'easy' rhymes means that they are an ideal candidate for a case study in the examination of the frequency of certain kinds of rhymes within the J-corpus, which we will perform here using methods of digital textual analysis. What follows is therefore not an attempt to pass aesthetic judgement on Jehan's texts. For the sake of simplicity, we will refer to the forms in question as 'easy' rhymes, yet the aim of this section is simply to reach a clearer understanding of the idiosyncratic uses of such rhymes that unite the corpus, and in so doing to add credence to the claim that the texts share an author.

Our focus here will be on two specific types of 'easy' rhyme which appear to be particularly common in Jehan's texts, therefore giving us the largest possible dataset, namely words ending in *-ment* and *-ant*. On occasion, while he largely prefers to keep them separate, perhaps on etymological or visual grounds, Jehan occasionally rhymes these forms interchangeably with each other, as in the stanza below from *Romme*, underscoring the fact that for Jehan, by the time he wrote his poems at the beginning of the fourteenth century, and in keeping with broader linguistic trends, *an* and *en* had an identical pronunciation.

De leur biensfez estoit par tot le renon grant ;
Il orent de .ii. anz .i. moult tres bel enfant :
Il estoit valeton, si l'amerent forment.
Li preudon n'en ot plus en trestot son vivant. (*Romme*, vv. 13–16)

To reduce the influence of orthographic variation on the data, which could result from scribal rather than authorial preferences, I also include those rhymes where the final *t* of *-ant* is replaced by *s* or *z*, since this variation does not appear to be caused by variation in the underlying sound of the rhyme, as evidenced by the stanzas below, in which *-ant*, *-ans*, and *-anz* rhyme interchangeably.[73] It is nevertheless notable that on the whole these forms are indeed kept separate, again perhaps for visual or etymological reasons rather than auditory ones, with stanzas containing a mixture of forms only occurring relatively rarely.[74]

Celui qui fu preudons si fu tous merveillans
De ce que son voisin li aloit devisans.
Il hauça sa main destre, si s'en ala seignant
Et puis si li a dit : "Vous estes mescreant,"
(*Mescreant*, vv. 49–52)

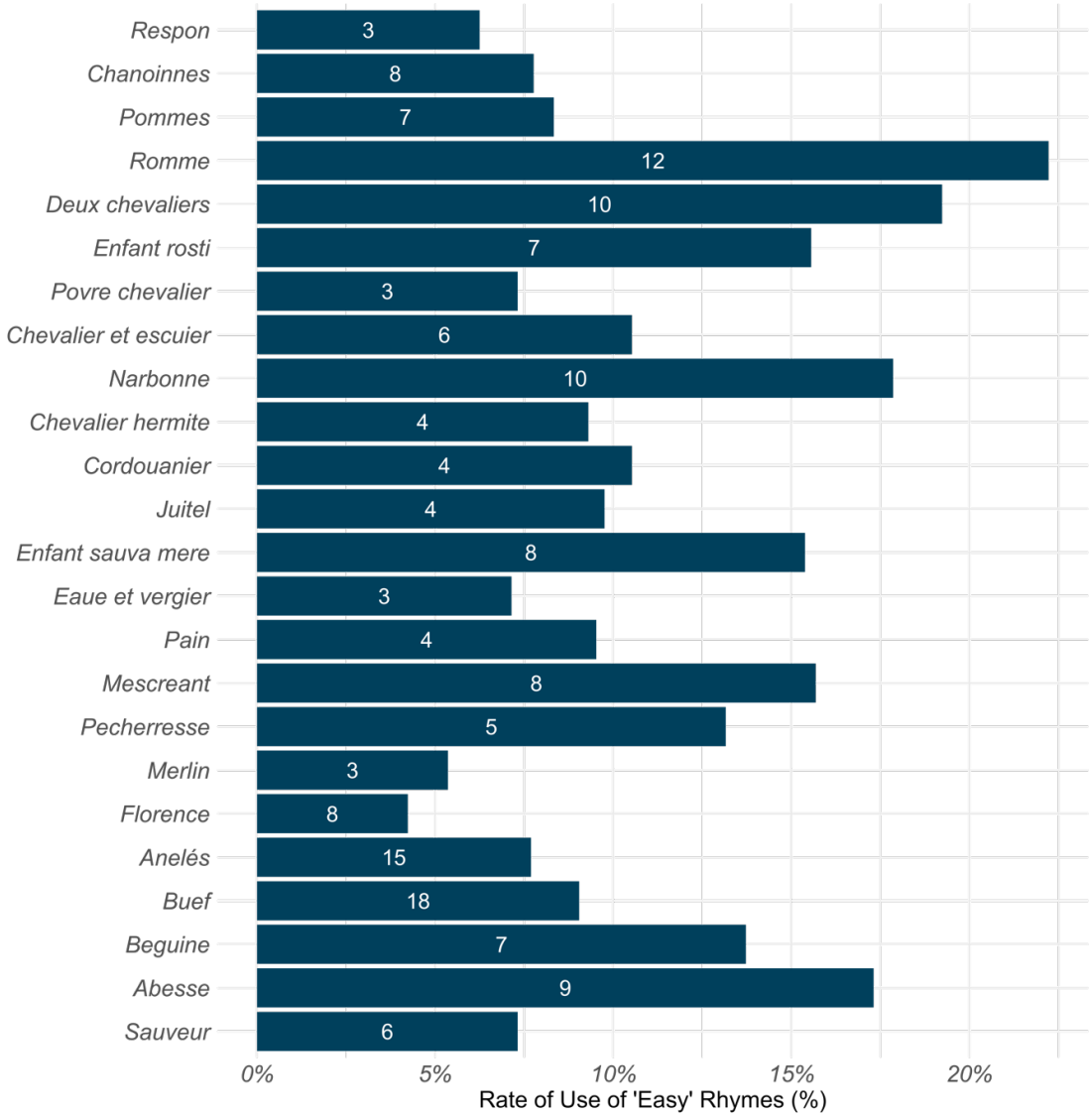
Li preudons respondi, moult liez et moult joians,
Pour ce que Diex ot fait vertuz suz ses enfans :
"Sire, a vostre gré faire sui touz obeissanz."
"Tu as droit," dit le pappe, "ja ni seras perdanz."
(*Buef*, vv. 769–772)

Through examination of the uses of 'easy' rhymes in Jehan's corpus of texts, as represented in Figure 3.57, we find a similar story to that which we observed for feminine rhymes, in that the use of such rhymes in the J-corpus is significantly variable. The data have a mean of a 10.1% usage of 'easy' rhymes, but rates for individual texts range from 4.2% (*Florence*) to 22.2% (*Romme*). It is difficult to come to any sort of conclusion as to why this variation would occur in the J-corpus, but there does appear to be a weak negative correlation between the number of stanzas contained within a text, and the rate at which that text contains 'easy' rhymes, that is to say, the shorter a text is, the more likely it is to make use of 'easy' rhymes

[73] The same does not appear to be true for *-mens* and *-menz*, which never rhyme with *-ment* in Jehan's texts. The regular expression which at least one line-final word in a given stanza has to match in order to be counted is therefore `an[tsz]$|ment$`.
[74] This would suggest that while, as we have seen, Jehan's texts do appear to have been destined for performance, there was nevertheless also a preference for some degree of aesthetic uniformity in the written medium, whether that arose from a preference by the author or by subsequent scribes.

(-0.37247). A Pearson's product-moment correlation test suggests that the correlation is not quite strong enough to determine statistical significance by our 95% threshold (p = 0.07306), but the very possibility of such a correlation is intriguing, as it could suggest that, if indeed we assume that such 'easy' rhymes require less artistic effort than others, then the works which demanded Jehan's greatest commitment to versificatory complexity were his longer texts.[75]

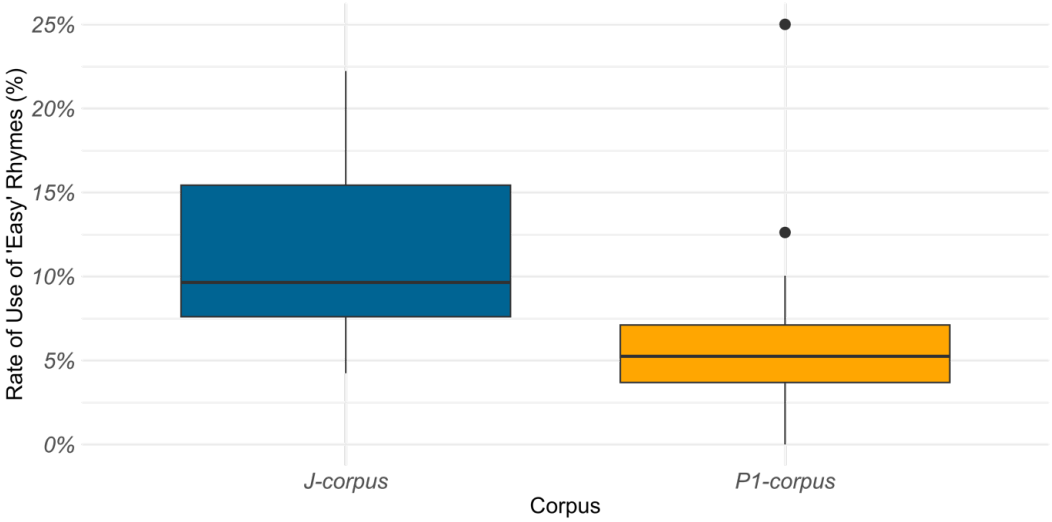
Figure 3.57 – rates of the use (x-axis) and raw numbers (figures in white) of stanzas containing 'easy' rhymes in the J-corpus.



[75] No such significant correlation appears between the use of feminine rhymes and text length.

A similarly noteworthy observation appears if we compare the data here to the rates of feminine rhymes in Figure 3.52, since we observe a weak negative correlation (-0.32074) between the rate of use of 'easy' rhymes, and the rate of use of feminine rhymes. That is, the more 'easy' rhymes Jehan uses in a given text, the fewer feminine rhymes he uses. Again, a p-value test results in a fail result by a small margin ($p = 0.1265$), but such a correlation would not be surprising, since a greater proportion of 'easy' rhymes, themselves masculine, would reduce the possibility of the appearance of feminine rhymes, and *vice versa*.

Figure 3.58 – box and whisker plots for the rate of use of 'easy' rhymes in the J- and P1-corpora.



When we compare the rates of use of 'easy' rhymes in the J-corpus to the P1-corpus, containing the poems found across fr24432 which are not usually attributed to Jehan, as represented in Figure 3.58, the results are striking.[76] Jehan appears on the whole to make much greater use of 'easy' rhymes when compared to other authors whose poems appear in the manuscript, with a median rate of 9.6% in his texts as opposed to 5.2% elsewhere. That said, the J-corpus arguably displays a greater variety of use than the P1-corpus, particularly in

[76] We recall that the P1-corpus excludes two fragmentary poems, *Lai d'amour 1* and *Trois mortes*. The former text makes use of 'easy' rhymes at a rate of 5.1%, almost exactly the median for this corpus, while the latter text does not use any 'easy' rhymes at all. Given the non-stanzaic form of most of the poems in the P1-corpus, the analysis of this corpus focusses on whether an individual line, rather than a full stanza, contains a so-called 'easy' rhyme.

the middle 50% of its data, displaying an inter-quartile range of 7.8% compared to just 3.4% for the P1-corpus. A Welch Two Sample t-test here is decisive ($p = 0.00001$), meaning that we are able to clearly distinguish Jehan's texts from those found elsewhere in fr24432 on the basis of their use of so-called 'easy' rhymes.

While the majority of texts in the P1-corpus therefore make lesser use of 'easy' rhymes than in the J-corpus, the presence of two significant upper outliers is significant. The greatest of these, with a higher rate of use of 'easy' rhymes even than any text of the J-corpus, is *Despensier*, a very short poem in octosyllabic rhyming couplets which, in its 40 lines, uses forms which fit our definition of 'easy' rhymes no less than 25% of the time. Explanation for this comes from the fact that the final word of each line contains the syllable *pen*, which in 24 of the lines appears within a form or derivation of the verb *penser*. Verb forms therefore predominate in the rhyming position, in particular present participles ending in *ant*, going some way to explain the elevated frequency of use of the so-called 'easy' rhymes in this text.

The second upper outlier within the P1-corpus is *Nostre Dame*, a tripartite text in praise of the Blessed Virgin which also happens to contain the lowest rate of feminine rhymes anywhere in fr24432 (12.6%). With a rate of use of 'easy' rhymes coincidentally also at 12.6%, the text is well within the range of Jehan's texts, but the fact that it stands out within the rest of the P1-corpus is a clear indication of the extent to which the J-corpus is distinct for this metric.

Given this additional discovery in *Nostre Dame* of a text with an unusually high use of 'easy' rhymes and an unusually low rate of feminine rhymes, we may once again interrogate, this time for the P1-corpus, the extent to which these two factors are correlated. As with the J-corpus, we observe a weak negative correlation between the use of feminine rhymes and the use of 'easy' rhymes (-0.40139), such that increased use of feminine rhymes is likely to be accompanied by decreased use of 'easy' rhymes. On this occasion, the calculated p-value is

statistically significant (0.00289), adding greater credence to the hypothesis reached for Jehan's texts. The fact that there is no great difference in the correlation coefficient between the J-corpus and the P1-corpus also supports our suggestion above that this phenomenon can be explained, at least in part, by the fact that increased use of the 'easy' rhymes in our class, which are masculine, leaves less space for the use of feminine rhymes.

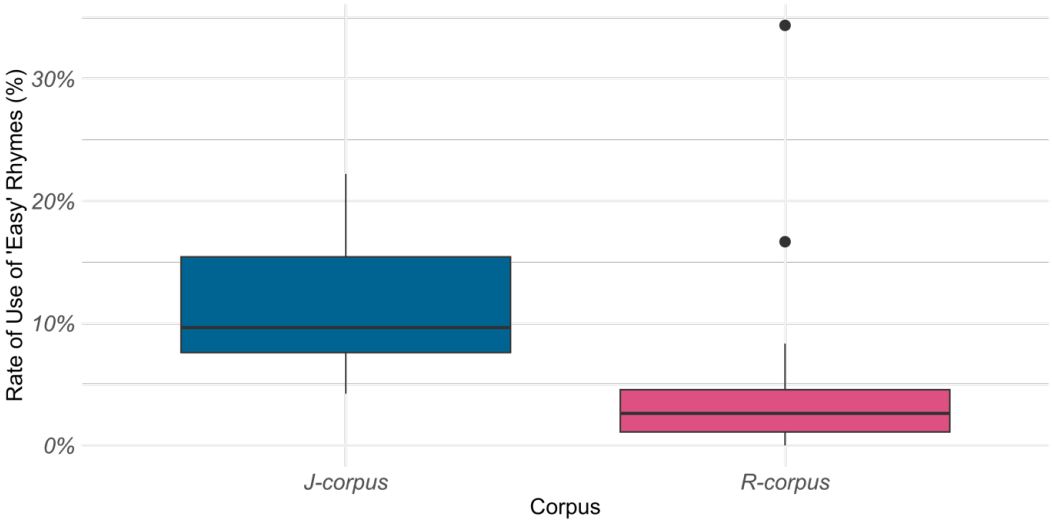
Turning to the lower end of the range of 'easy' rhymes observed in the P1-corpus, we note that while there are no statistical outliers, there are nevertheless a number of texts which are worthy of attention. The most extreme example of this is *Complainte*, a text by Rutebeuf, which makes no use of 'easy' rhymes anywhere in its 144 lines. All in all, this seems to fit with the rates of use observed for other of Rutebeuf's texts contained within the P1-corpus, certainly when compared to Jehan's texts, further supporting our suggestion that the use of such rhymes can be used as a means of indicating relative unity in a way that can support the attribution of authorship. *Vie du monde* displays the highest rate of Rutebeuf's texts within the manuscript (9.1%), followed by *Repentance* (8.3%), and *Charlot* (4.8%). *Erberie*, which has been excluded from the P1-corpus on account of its mixed verse and prose form, nevertheless offers a useful intermediate point between these texts and *Complainte*, since in its 111 verse lines, it contains just 2 rhymes which satisfy our classification as 'easy' rhymes, giving a rate of 1.8%.

On the basis of these observations, it is worthwhile comparing the J-corpus to the R-corpus containing the complete surviving poetic works of Rutebeuf, in order to determine whether it is possible to identify a distinct use of 'easy' rhymes between the texts of the J-corpus and those of another author for whom attribution is more assured.[77] The data are presented in

[77] Given the formal variety of Rutebeuf's *œuvre*, when calculating the rate of 'easy' rhymes for the R-corpus, it is necessary to consider the texts on a line-by-line basis, as with the P1-corpus, since we are often not able to divide the text into monorhyme stanzas. We note also that the texts attributed to Rutebeuf which appear in fr24432 are not exactly as they appear in the edition of the texts we use for the R-corpus (Bastin & Faral 1959–1960), and so it is possible for divergences to appear between the rates used in the same texts in the P1- and R-corpora.

Figure 3.59. Again, for the R-corpus, we note that the median use of 'easy' rhymes is much lower than that observed for the J-corpus, with 13 texts of the 56 in the R-corpus not containing a single instance of an 'easy' rhyme. The presence of the two upper outliers adds somewhat to the sense that the work of a single author need not be consistent in its use of 'easy' rhymes, and in this particular case adds to the sense of aesthetic variety that is present within Rutebeuf's *œuvre*. The highest rate anywhere in the R-corpus, displayed in *La chanson de Puille* (34.4%), appears to result from the aesthetic concept of the poem, with 'easy' rhymes appearing in the fourth and eighth lines of each of the poem's 8-line stanzas. The relatively high rate in *De Brichemer* (16.7%), meanwhile, appears more straightforwardly to be a function of the brevity of this poem, with 'easy' rhymes appearing in just 4 of the poem's 24 lines, all rhyming with each other as one of the two rhymes in an 8-line alternating stanza.

Figure 3.59 – box and whisker plots for the rate of use of 'easy' rhymes in the J- and R-corpora.



On the whole, then, the R-corpus contains significantly lower rates than the J-corpus ($p = 0.00000$), just as was the case for the P1-corpus, and we can say with some certainty that Jehan's use of so-called 'easy' rhymes distinguishes the J-corpus from other texts.

As a final point, returning our attention to *Guillaume* and *Robert*, we also note that both display rates of use of 'easy' rhymes which are within the range of the J-corpus (10.1% and 6.7%

respectively), and in the case of *Guillaume* in particular, this figure is on the edge of being considered an outlier for the P1-corpus, while only being slightly above the median of the J-corpus. As a result, this method of analysis allows us to once again strengthen our attribution of the J-corpus to Jehan, while also reinforcing the possibility that *Guillaume* and *Robert* merit inclusion in his *œuvre*, too.

3.6.3 Word frequency in rhyming position

Munk Olsen (1963b: 160) suggests a further element of versification which may prove effective in comparing Jehan's poems. It may be fruitful, Munk Olsen argues, to consider *associations des rimes*, that is, which words appear most frequently in rhyming position when a certain rhyme is required. While he does not perform a full analysis, Munk Olsen's method appears promising, since it allows him to compare his corpus of Jehan's texts to the alexandrine version of the *Vie de saint Alexis* through analysis of the rhymes *-on(s)* and *-is*.[78] In doing so, Munk Olsen observes differences in the words most frequently found for each rhyme, and reinforces his claim that the text cannot be said to be another work by Jehan.

Digital textual analysis, however, allows us to go one step further, looking not simply at which words appear most frequently for a given rhyme, but at which words are more likely to appear in rhyming position than at other moments during the text. To do so, we calculate the rate at which a given word appears in rhyming position, as well as the rate at which that same word appears across all remaining tokens of a given text.[79] We then divide the former rate by the latter to determine what I term the *rhyme coefficient*, that is, how much more likely a given word is to appear in rhyming position than elsewhere in the line. This method allows us to

[78] The corpus Munk Olsen uses for this analysis is shorter than the 1978 corpus on which he eventually settles.
[79] The line-final rate is calculated by dividing the number of observed instances of a given lemma in final position by the total number of line-final tokens in the respective corpus (6,829 in the J-corpus, 12,668 in the R-corpus, and 27,338 in the P1-corpus), while the non-final rate is calculated using the total number of non-final tokens in the corpus (52,636, 64,647, and 138,209).

make a critical distinction between those words which appear frequently due to the simple fact that they are often used across the corpus, as a result of authorial preference or narrative necessity, and those words which could be described as 'stock' rhyming words, appearing when necessary to complete a rhyme, but rarely used elsewhere.

For this section of the analysis, we do not focus on the particular word forms that appear at the end of lines, rather their associated lemmata, that is, the arbitrarily determined 'dictionary' form of a lexeme, as tagged by RNNTagger. This is an important distinction since it allows us to ignore the influence of both orthographic and inflectional variation and focus on the underlying lexeme choice. Rates of appearance are therefore calculated based on the appearance of lemmata, regardless of the exact word form to which they are associated. To avoid cases in which the data would be skewed by the appearance of only a few forms of a word, we limit our examinations here to lemmata which appear on at least 10 occasions anywhere in the corpus, and at least once in rhyming position.

As an initial proof of concept, let us apply this method to the R-corpus, for which the single-author hypothesis is more assured. Across the 56 texts of the R-corpus, we observe some 49 lemmata which match our selection criteria and which appear with a rhyme coefficient greater than 1, meaning that they are more commonly used in rhyming position than elsewhere in a line. Of these, *longuement* (681.3), *jugement* (264.9), *mentir* (100.9), *errer* (75.6), *durement* (63.1), *semblant* (56.8), *neant* (53.4), and *vent* (50.5) are at least 50 times more likely to be included at the end of a line, suggesting that these are words which Rutebeuf significantly favoured in rhyming position, and indicating that this approach has significant validity as a means of identifying the individual versificatory 'fingerprint' of a given author.

Applying this analysis across Jehan's texts, we also find a number of trends of note. Table 3.29 shows the top 20 lemmata across the corpus, when ordered by descending rhyme

coefficient. Firstly, we discover that forms of *certainement*, *envie*, and *demeuree* only ever appear in Jehan's texts in rhyming position, marking these forms out as words which may have been retained by Jehan for use solely in such a position. We also note that many of these lemmata with a high rhyme coefficient, including *certainement*, *vraiment*, *commandement*, and *sauvement*, count as examples of the so-called 'easy' rhymes discussed above.[80]

Table 3.29 – counts and rates of lemmata which appear on more than 10 occasions in the J-corpus, limited to the top 20 results.

Lemma	Line-final count	Line-final rate	Non-final count	Non-final rate	Rhyme coefficient
certainement	13	0.19%	0	0.00%	—
envie	11	0.16%	0	0.00%	—
demeuree	10	0.15%	0	0.00%	—
contree	12	0.18%	1	0.00%	92.5
plevir	12	0.18%	1	0.00%	92.5
vraiment	12	0.18%	1	0.00%	92.5
arreter	11	0.16%	1	0.00%	84.8
partie	21	0.31%	2	0.00%	80.9
commandement	10	0.15%	1	0.00%	77.1
appareiller	9	0.13%	1	0.00%	69.4
usage	9	0.13%	1	0.00%	69.4
tenter	13	0.19%	2	0.00%	50.1
pensee	19	0.28%	3	0.01%	48.8
endurer	11	0.16%	2	0.00%	42.4
chiere	10	0.15%	2	0.00%	38.5
corage	20	0.29%	4	0.01%	38.5
esbahir	9	0.13%	2	0.00%	34.7
debonnaire	8	0.12%	2	0.00%	30.8
matiere	8	0.12%	2	0.00%	30.8
sauvement	8	0.12%	2	0.00%	30.8

[80] Across all lemmata which appear in rhyming position at least once and in the whole corpus on more than 10 occasions, we find a further 20 such lemmata with a rhyme coefficient greater than 1: *largement* (23.1), *jugement* (17.3), *isnellement* (14.3), *joiant* (12.3), *tourment* (12.3), *talent* (12.0), *malement* (7.7), *longuement* (6.9), *couvent* (6.4), *erramment* (5.8), *avant* (4.2), *briément* (3.5), *argent* (3.3), *maintenant* (3.1), *gent* (1.9), *dolent* (1.9), *souvent* (1.9), *doucement* (1.7), *enfant* (1.4), and *fortement* (1.3). Given that this is also the case in the texts of the R-corpus, with 15 of the 49 forms with a rhyme coefficient above 1 counting as 'easy' rhymes, we are able to further resist the claims of previous scholarship that an abundance of such forms has anything to do with poor versificatory practice. Rather, a significant use of so-called 'easy' rhymes may simply be evidence that a given author prefers to use such forms in rhyming position rather than at any other point in a line.

Based on these examples of words with high rhyme coefficients, Jehan therefore appears to have had a stock of nouns and adverbs which he used largely in rhyming position when required to complete a stanza.[81] This is in addition, of course, to the many verbal forms satisfying the criteria for 'easy' rhymes which contribute to the overall trends observed above, and yet notably, none of the verbs with the highest rhyme coefficients appear with particular frequency in the present participle form, for example, suggesting that in forming 'easy' rhymes, Jehan may have extracted his forms from whatever verb was necessary to the narrative, rather than having a stock of verbs which he preferred to use in such situations.

Evidence from both the J- and R-corpora therefore suggests that the rates at which certain words appear in rhyming position is as a direct result of authorial preference. By also comparing such forms with the highest rhyme coefficients to the P1-corpus, as a broader reference corpus containing the work of a number of authors, it is possible to identify trends which assist us in attributing the authorship of the J-corpus to a single author. Table 3.30 displays the top 20 lemmata for the P1-corpus, when ordered by rhyme coefficient, allowing us to make direct comparisons to the texts of the J-corpus.

Table 3.30 – counts and rates of lemmata which appear on more than 10 occasions in P1-corpus, limited to the top 20 results.

Lemma	Line-final count	Line-final rate	Non-final count	Non-final rate	Rhyme coefficient
seigneurie	15	0.05%	0	0.00%	—
baudeur	13	0.05%	0	0.00%	—
desconforter	11	0.04%	0	0.00%	—
desesperance	11	0.04%	0	0.00%	—
devise	11	0.04%	0	0.00%	—
adrecier	10	0.04%	0	0.00%	—
agreer	10	0.04%	0	0.00%	—

[81] At the other end of the scale, words with rhyme coefficients close to zero are usually function words which occur with significant frequency across the whole text, but which are rare in rhyming position by dint of being forms which must usually be followed immediately by a more contentful word. A rhyme coefficient of less than one therefore most likely results from linguistic factors, rather than being the result of an authorial choice to include the word less frequently in rhyming position. Our attention here should thus be squarely focussed on the forms with the highest rhyme coefficients.

atorner	10	0.04%	0	0.00%	—
misericorde	10	0.04%	0	0.00%	—
signifiance	10	0.04%	0	0.00%	—
afaire	41	0.15%	1	0.00%	207.3
esjouir	21	0.08%	1	0.00%	106.2
apert	17	0.06%	1	0.00%	85.9
chevalerie	15	0.05%	1	0.00%	75.8
entierement	15	0.05%	1	0.00%	75.8
venue	15	0.05%	1	0.00%	75.8
lignee	14	0.05%	1	0.00%	70.8
apertement	13	0.05%	1	0.00%	65.7
baillir	12	0.04%	1	0.00%	60.7
martire	12	0.04%	1	0.00%	60.7

The first main observation of note is that none of the lemmata which appear exclusively in rhyming position in the P1-corpus are the same as those which only appear in rhyming position in the J-corpus. Looking more broadly, if we compare the rhyme coefficients of the reference corpus to the highest rhyme coefficients in Jehan's texts for individual lemmata, we can identify a number of cases in which the usage appears to be significantly different between the two corpora, namely *vraiment* (J-corpus: 92.5; P1-corpus: 8.7), *usage* (69.4 to 7.1), *pensee* (48.8 to 5.5), and *visage* (30.8 to 3.4), all of which have much higher rhyme coefficients in the J-corpus than in the P1-corpus. Looking in the other direction is similarly fruitful, with *jornee* (1.9 to 27.8), *dependre* (2.4 to 30.3), *defendre* (2.8 to 24.4), and *guise* (1.1 to 9.7) all appearing with markedly greater rhyme coefficients in the reference corpus than in Jehan's texts.[82]

It is therefore clear that, for certain words, the rhyme coefficient differs a great deal between the J- and P1-corpora, suggesting a role played by authorial preferences and therefore supporting the single-author hypothesis for the J-corpus. Overall, however, the picture is not quite as clear, since despite the examples we have seen, there is still a correlation between the rhyme coefficients observed in the J-corpus and those observed for poems in the P1-corpus.

[82] The R-corpus also overlaps with the J-corpus somewhat, sharing some 28 terms with a rhyme coefficient of greater than 1, with varying levels of similarity in the value of the rhyme coefficient.

Examining those lemmata within the J-corpus which appear at the rhyme in at least one instance and at least 10 times across the corpus, and also appear in the P1-corpus, we find a statistically significant, strongly positive correlation (0.59336, p = 0) between the rhyme coefficients for individual lemmata. We can say with some certainty, then, that with the exception of a few examples, the lemmata which are more likely to appear at the rhyme in the J-corpus are also those which are more likely to do so in the P1-corpus.

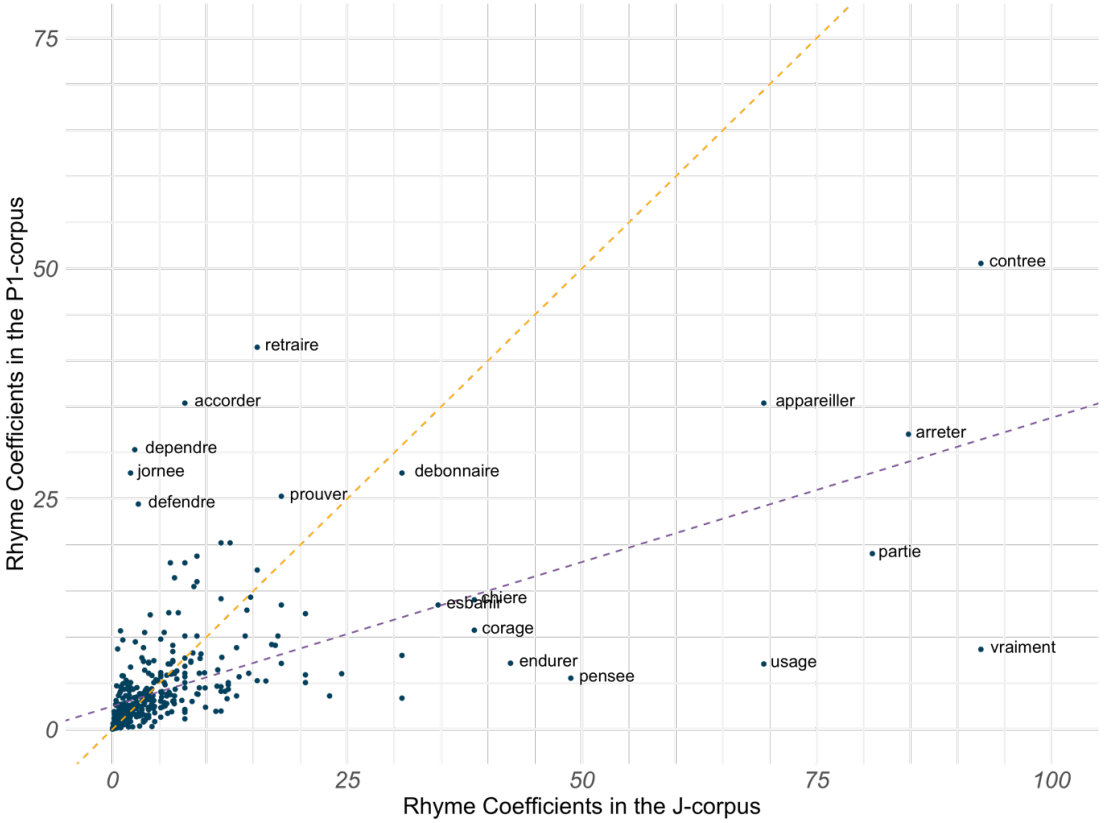
However, by producing a linear regression model for these data, a so-called line of best fit, we are able to discover that the link between the two corpora is not one-to-one, in that it is not simply the case that lemmata tend to have the same rhyme coefficient across both corpora. Instead, a calculated regression line (intercept = 2.50, gradient = 0.31291) suggests that a lemma in the J-corpus with a rhyme coefficient of 100 would be expected to appear with a rhyme coefficient of just 33.79 in the P1-corpus. While we can therefore expect certain lemmata to have a greater rhyme coefficient across both the J- and P1- corpora, Jehan reserves certain words for use solely, or primarily, at the ends of lines to a much greater extent than those authors whose work his poems appear alongside.

Figure 3.60 maps rhyme coefficients in both corpora as a scatter plot, along with the calculated regression line (in purple), and a line indicating a putative one-to-one correlation (in orange).[83] From it, we observe that there are indeed certain words which do share something resembling a one-to-one connection between their rhyme coefficients in the J- and P1-corpora, including *prouver* and *debonnaire*. For these words, it seems, the rhyme coefficient may be relatively predetermined regardless of authorial identity, perhaps resulting from structural, linguistic factors determining where the term tends to fall, or indeed also from

[83] A selection of points have been labelled with their corresponding lemmata, but most have not, to ease readability. Lemmata have been excluded where the rhyme coefficient is infinite in either corpus, that is, where the only examples of a lemma are in the rhyming position and no examples are to be found elsewhere.

broader contemporary expectations of style. Alongside this, we also get a clearer picture of the outlier status of lemmata such as *endurer*, *pensee*, *usage*, and *vraiment*, which have much higher rhyme coefficients in the J-corpus than in the P1-corpus, as well as *defendre*, *jornee*, *dependre*, *accorder*, and *retraire*, for which the opposite is true. Thus, while the rhyme coefficients for some lemmata appear to be consistent across the work of different poets, or at least roughly in line with a consistent ratio of usage, there are still several lemmata for which it appears that authorial identity is the most significant factor in determining their rhyme coefficient.[84]

Figure 3.60 – scatter plot for the rhyme coefficients of lemmata found in the rhyming position of both the texts of the J- and P1-corpora.



[84] The methods discussed in this sub-section lend themselves best to examination of whole corpora of texts, rather than individual texts, since the latter are likely to be influenced in their word choices by their content. It has therefore not been possible to apply this analysis to *Guillaume* and *Robert* individually.

3.6.4 Admixtures

A final avenue of enquiry regarding Jehan's versificatory practices is that of admixtures, that is, cases where we observe an increased co-occurrence of one particular word in rhyming position when another word is also present in another rhyming position in the same stanza. Again, we look here at lemmata rather than word forms, so as to eliminate the possible impacts of orthographic and inflectional variation. Seeking those examples which are likely to be most assured in the trends they represent, our threshold for this analysis is to look only at lemmata which appear in at the rhyme on 5 or more occasions in the J-corpus, since the certainty with which we can say that two lemmata are connected at the rhyme decreases the lower the frequency of those words.

To take an extreme example of why such a threshold is necessary, in the following stanza we find three words, 'ordure', 'ardure', and 'vesteüre', which only appear at the rhyme once in the corpus of Jehan's texts. It would be an error, therefore, to assume that their admixing here means that such forms would only ever appear together, since there are insufficient examples to be able to draw such a conclusion for certain.

> En escommenjement vesqui et en ordure ;
> De li faire rasoudre n'avoit talent ne cure.
> Ens el moustier entra courouciez en ardure ;
> Le prestre pour chanter ot pris sa vesteüre. (*Chanoinnes*, vv. 101–104)

Having isolated only those lemmata which are deemed to exist in sufficient numbers to reliably suggest the presence of an admixture, we can then calculate the ratio between the number of forms of a lemma appearing in the context of another lemma to those which do not. Thus, a lemma which appears 5 times alongside a context lemma and 2 times elsewhere has a ratio of 2.5(:1). A lemma is said to exist in a prominent admixture with another lemma

if this ratio is greater than one, meaning that the lemma appears alongside its associated context lemma more frequently than it does elsewhere.[85]

There are 34 pairs of lemmata which satisfy these criteria within Jehan's texts, suggesting that they are pairings which the poet chose to use together on a particularly frequent basis. The lemma *faire* has four associated lemmata which appear in prominent admixture with it (*debonnaire* (ratio = 7), *retraire* (7), *contraire* (5), and *exemplaire* (5)),[86] all of which, we note, are used when *faire* is in its infinitive form. The same is not true of *avoir*, whose two associated lemmata (*voloir*, 1.4; *remanoir*, 1.2) can be used in a number of forms, whereas in the case of *dire* (*martire*, —; *despit*, 5; *Jesus-Christ*, 5; *respit*, 2.5), the necessary forms for the rhyme are the third person singular in the present tense, or the past participle, for the last three forms, and the infinitive for the first form.

The case of *martire* is particularly notable in that this lemma does not appear in any rhyming position other than alongside *dire*, meaning that the ratio is infinite, and suggesting a strong association between these two words. The same is true of *Rome*, which only ever appears alongside forms of *somme*. In this case, we also find that the same connection exists in the opposite direction, with *somme* (5) appearing almost exclusively in the presence of *Rome*. Such bivalent pairs can be found elsewhere for *debonnaire* (1.67) and *retraire* (1.67), *chemin* (2) and *fin* (1.33), *maniere* (2) and *chiere* (1.5), and *mere* (1.75) and *pere* (7.5). In some cases, as between *mere* and *pere*, there is a clear directionality to this bivalence, with the higher ratio associated with *pere* suggesting that *pere* is triggered by the presence of *mere*, rather than the other way around, whereas in other cases the pairs appear to be more mutually reinforcing.

[85] It is likely that word pairings could indeed be prominent admixtures with a ratio of less than 1, if a given context lemma is unusual. Looking only at forms which are as likely, or indeed more likely, to appear in the context of a given lemma than anywhere else, however, allows us to reach more assured conclusions, since the strength of the admixture is likely to be particularly strong.
[86] We say that *faire* is therefore a *context lemma* for these lemmata.

In the case of *mere* and *pere*, as well as forming a bivalent pair, the two are also both context lemmata for *frere* (6 with *mere*, 2.5 with *pere*), while *mere* also takes *amer* (1.75) as a context lemma and is itself a context lemma for *clair* (5). Of the remaining examples, we find four lemmata with *vie* as a context lemma (*compaignie*, 1.66; *folie*, 1.29; *envie*, 1.2; *Marie*, 1.13), four with *enfant* (*grant*, 2.2; *joiant*, 1.67; *avant*, 1.17; *maintenant*, 1.17), two with *certain* (*plein*, 6; *vilain*, 2) and *corage* (*usage*, 8; *image*, 2), and one each with *Marie* (*vilenie*, 4), *mie* (*compaignie*, 1.18), and *raison* (*maison*, 1.13). Interestingly, none of these lemmata are those which we identified as having a particularly high rhyme coefficient, thus reinforcing the suggestion that Jehan chose these forms for the rhyming position on the bases of the surrounding rhyming lemmata, rather than simply because he favoured the word in rhyming position.

Applying this approach to the other poetic texts in fr24432, it is necessary to limit our scope to other texts in monorhyme quatrains, so as to ensure that co-occurrence at the end of a line in the same stanza does in fact correspond to two words sharing a rhyme. Given our previous assertions that *Guillaume* and *Robert* could possibly be the work of Jehan, we exclude them from the P1a-corpus for this analysis, resulting in a grouping of 5 texts we shall term the P1b-corpus.[87] At just 1,703 lines, the P1b-corpus is too small to reach many firm comparisons to the use of admixtures found in the J-corpus, simply since there are often not enough examples to confirm the validity of a particular observation. That said, we are able to determine that Jehan does use *debonnaire* alongside *faire* more frequently than in the P1b-corpus (ratio in J-corpus: 7, ratio in P1b-corpus: 0.33, $p = 0.046912$), contributing to our more general sense that the texts of the J-corpus display a number of versificatory tendencies that are distinct from those of other poets, and reinforcing the strength of the attribution of these texts to a single author.

[87] *Bible*, *Vision*, *Vie du Monde*, *Mais*, and *Menage*.

Turning our attention finally to *Guillaume* and *Robert*, again these texts are too short on their own to permit a fine-grain analysis of the admixtures that appear within them, but it is nevertheless notable that both texts contain 5 of the same admixture pairs as are also found in the J-corpus, namely *faire–debonnaire*, *dire–respit*, *mie–compagnie*, and two bivalent pairs *mere–pere* and *chiere–manière*, while *Guillaume* displays a further 3 pairs (*dire–martire*, *dire–Jesus-Christ*, and *enfant–joiant*), and *Robert* a further 13 (*vie* with *compagnie*, *folie*, and *Marie*; *mere* with *frere* and *clair*; *enfant* with *grant* and *avant*; *certain* with *plein* and *vilain*; *dire–despit*; *raison–maison*; *amer–mere*, and the bi-valent pair *Romme–somme*). In all cases, the admixture ratios calculated for these pairs, when compared with those of the J-corpus, result in p-values all a fair deal greater than 0.05, thus indicating that at the very least there is no evidence that any of these admixture pairs are used differently in *Guillaume* and *Robert* to in the J-corpus. While the utility of this method is therefore limited when used on these short texts, it nevertheless still allows us to strengthen the growing body of evidence that these texts ought to be attributed to Jehan. From a methodological perspective, meanwhile, there is an indication that this method may be even more usefully applicable for questions of authorship concerning longer poetic texts.

To conclude this section, we have seen that while historical claims in the scholarship as to the quality of Jehan's poetry are largely unhelpful, they nevertheless point to useful avenues through which his versification can be considered by means of digital poetic analysis. Often, Jehan's tendencies come close to those observed in the poems found elsewhere in fr24432, and yet for each of the metrics considered, the texts attributed to Jehan have been shown to be qualitatively and quantitatively distinct from them such that, despite their individual diversity, we may nevertheless still assert that a digital analysis of the versificatory practices of the J-corpus contributes to an overarching sense of unity which allows us to support the single-author hypothesis.

3.7 Shared vocabulary and structures

3.7.1 Existing observations

Building on observations made in passing by Naetebus (1891: 62, 75), the evidence which Morawski (1939, 1940) considers amongst the most significant in allowing the collective attribution of the *dits* to Jehan is the presence of structures which are shared across the texts of the corpus. The most striking of these is the observation that *Povre chevalier* shares a full stanza with *Abesse*, one of the two texts in Morawski's corpus found outside of fr24432.

Por ce que Nostre Dame scet paier largement,
Veil ge metre m'entente a moustrer plainement
Les miracles qu'el fait a tous communement
Qui la veulent servir de cuer parfaitement
(*Povre chevalier*, vv. 5–8)

Pour ce que Nostre Dame scet paier largement,
Weil je metre m'entente a moustrer plainnement
Les miracles qu'el fait a tous communaument
Qui la weilent servir de cuer parfaitement
(*Abesse*, vv. 9–12)

Morawski observes that many of the structures across Jehan's poems relate to shared elements of narrative in the story,[88] but also that there is a particularly significant selection of shared structures related to the texts' religious contents. God, for example, is described as 'le roi de maesté' in a total of three separate texts,[89] while Mary appears as the one 'qui le cors Dieu porta' in no fewer than five texts in the corpus.[90]

However, while these observations certainly add to the sense of unity we have already observed across the corpus, Munk Olsen (1963b: 152) appears justified in his criticism of treating such features as objective indicators of shared authorship, particularly for religious language which, by its very nature, is likely to rely on clichés and commonplaces. Outside of such examples, to claim that shared structures are proof of shared ownership also still relies on the assumption that an author's work is always written within a limited time period, with

[88] For example, 'Tost et isnelement entra en la chapelle' (*Povre chevalier*, v. 90; *Chevalier et escuier*, v. 107).
[89] *Respon* (vv. 90, 176): 'le roy de maesté'; *Pommes* (v. 306): 'le roy de maisté'; *Abesse* (v. 131): 'le roy de majesté'.
[90] *Respon* (v. 121), *Romme* (v. 213), *Deux chevaliers* (v. 113), *Chevalier et escuier* (v. 78): 'qui le cors Dieu porta'; *Narbonne* (v. 100): 'qui le corps Dieu porta'. Morawski appears to miss the structure found in *Romme*.

easy access to past works, a view which treats the author like 'un cerveau électronique dans lequel une impulsion de contenu donné déclenche automatiquement et infailliblement l'impulsion d'expression correspondante,' in other words a machine which, in consistent circumstances, will be incapable of varying its output.

Even the stanza shared between *Povre chevalier* and *Abesse*, while encouraging as part of a broader case for shared authorship, is not enough on its own to generate a convincing argument, since it is also equally plausible that the author of one or other of these texts borrowed this relatively formulaic introductory section that they had read in the other. Indeed, the inclusion could even have been made by a later scribe familiar with both texts, since the stanza is discrete and, in the case of *Abesse*, it is not a required element of the poem from either a grammatical or a narrative perspective, since the exposition of the narrator's intention to tell his story has already been announced in earlier lines. All of this is not to say that the observation of shared structures is necessarily meaningless to debates regarding authorship, only that we cannot see them, as Morawski arguably did, as a silver bullet which can be used to decisively prove shared authorship.

Moving beyond this primary focus on the simple observation of the same structures in more than one text, Kleist (1973: 44–45) builds on a statement made in passing by Morawski (1939: 332), that 'ce sont toujours les mêmes expressions, formules et rimes qui reviennent en pareille occurrence.' For Kleist, this is a particularly significant observation, that not only are there recurring structures within Jehan's texts, but that these recurrences appear within comparable narrative contexts. Examining questions of repetition from this perspective, Kleist argues, allows us to differentiate between conventional repetition and that which may genuinely be an indicator of shared authorship.

To illustrate his point, Kleist considers the case of a shared narrative segment in two texts attributed to Jehan which appear consecutively in fr24432, *Povre chevalier* and *Chevalier et escuier*. Both stories share three narrative phases. They tell of a character who, in a moment of desperation, sees a chapel and decides to go inside, in the former case a lady whose husband has promised her to the Devil, and in the latter case a knight who has fallen from grace (phase 1). This character turns their horse towards the chapel (phase 2), and then enters the chapel and goes towards the image of Mary (phase 3). While far from the most original literary motif, it is certainly notable that such a narrative progression is shared between both poems, especially since they are adjacent within fr24432. Beyond the identical line already identified by Morawski and given above in note 88, which takes place during phase 3, Kleist also notes that the stanzas containing the narrative phases 2 and 3 share a rhyme in both texts, and argues that, in this case 'handelt es sich nicht bloß um zufällige Übereinstimmungen, sondern um ein Stilprinzip, das man als Technik der Versatzstücke bezeichnen könnte' (1973: 46).

Some of Kleist's evidence may be described as tentative, particularly the shared rhymes in phases 2 and 3 which contain 'torna' and 'chapelle,' both of which appear especially likely to appear in this particular narrative context. There is also the possible argument that narrative similarities such as these were the primary reason for the juxtaposition of both texts in fr24432, regardless of whether or not they share an author. Further, we should note that these motifs are not unique to these two texts, and instead exist within established textual traditions. *Povre chevalier* corresponds to Tubach's (1969) types 5115, *Virgin, Blessed, comes to devil instead of his victim* and 5283, *Wife pledged to devil*, while *Chevalier et escuier* appears close to types 3570, *Pact with Devil V*, and 5133, *Virgin, Blessed, renunciation of refused*. All these text types are well attested, such that the narrative phases and their corresponding shared structures identified by Kleist cannot be said to be unique to these two texts.

There are thus reasons to be cautious about this evidence, yet the combination of this evidence with other similarities, not to mention the texts' codicological context, certainly reinforces the possibility that there was influence between these texts that was acknowledged by contemporary readers, whether or not that means they were written by the same individual. Through the digital and statistical analysis presented in this section, we will examine similar ways in which the single-author hypothesis can be strengthened, or not, by the presence of shared structures across the J-corpus. In doing so, we will seek to show that these points of narrative correspondence are not the only evidence of structural coherence displayed by its texts, allowing us to further progress the shared-author hypothesis.

3.7.2 Line-initial words

Munk Olsen (1963b: 155) argues, counter to the views of Morawski (1939, 1940) in particular, that there is a danger to solely considering repetitions which contain specific, narrative-relevant meaning, since these are more likely to be required by the subject matter than as a result of authorial preference. Instead, Munk Olsen argues, we should consider the use of more semantically 'empty' words, such as *et* and *quand*, in order to discover whether texts contain a consistent stylistic profile since, compared to content words which may be determined by narrative requirements, in the case of function words 'il est plus probable qu'ils représentent des "tics" chez l'auteur.'

There certainly appears to be value in such an approach, particularly when combined with methods containing a more holistic narrative focus, such as that of Kleist. Munk Olsen applies this method to some effect in his comparison of Jehan's *dits* to the alexandrine *Vie de saint Alexis* (1963b: 159), using it to further his claim that the text cannot be considered part of Jehan's *œuvre*, but his totals are not drawn from the full 1978 corpus of texts he assembled later, and which we have taken as the J-corpus, only the earlier corpus given by Morawski (1939). Perhaps for want of a means to automate what would otherwise be a drawn-out

process of counting the frequency with which each word appears at the beginning of a line, Munk Olsen also only offers totals for a few words, and his conclusions for the *Vie de saint Alexis* are therefore not driven by the whole data, and make no attempt to determine statistical significance. Munk Olsen apparently also largely limits his examination to function words which are in some ways more contentful,[91] thus going against his otherwise valid suggestion that it is the semantically 'empty' words which are of greatest interest for an analysis of authorial 'tics.' There is space, therefore, for a more data-driven approach to the same questions, which is what we shall attempt to perform here.

As in our work on line-final forms from page 189, our focus here is on machine-tagged lemmata, rather than observed word forms, since this approach allows us to bypass any scribal or inflectional variation. In the case of these data, where function words such as *le* and *la* are more common, it also enables us to distinguish between uses of these same word forms which can be both determiners and pronouns. Grammatical gender, number, and case distinctions between function words are all assumed here to be examples of inflectional rather than lexical variation, such that the determiners *le*, *la*, and *les* are subsumed under the single lemma *le*, while the pronominal forms *il*, *elle*, *le*, *li*, etc. are all assigned to the single lemma *il*.

Looking at the six most common lemmata in each of the texts of the J-corpus, presented in Table 3.31, the most striking feature is their variability.[92] Several lemmata that appear, such as *le*, *que*, and *et*, are consistently high in frequency, but there is no established order across all of the texts. Most of these words are, unsurprisingly, function words, but there are a few appearances of more contentful words, most notably of *mais* as the third most common lemma at the beginnings of lines in *Florence*, and *seignor* in sixth place in *Abesse*.

[91] *mais*, *car*, *lors*, *or*, *puis*, *adont*, *atant*, and *adoncques*, alongside the more empty words *et* and *quand*.
[92] The decision to look specifically at the 6 most frequent lemmata is largely an arbitrary one, but examination of the most frequent lemmata in particular is likely to yield the most reliable data.

Table 3.31 – the six most common lemmata at the beginning of a line for each of the texts in the J-corpus.

Text	1	2	3	4	5	6
Respon	le	il	de	que	a	et
Chanoinnes	le	et	a	que	quand	de
Pommes	le	que	quand	et	a	se
Romme	le	de	et	que	a	il
Deux chevaliers	le	que	de	et	a	lors
Enfant rosti	le	et	a	lors	son	de
Povre chevalier	que	le	de	a	car	en
Chevalier et escuier	et	le	je	que	quand	il
Narbonne	le	a	et	lors	quand	que
Chevalier hermite	le	et	que	quand	de	a
Cordouanier	le	et	de	a	en	or
Juitel	le	quand	et	que	a	car
Enfant sauva mere	le	et	que	a	de	quand
Eaue et vergier	le	que	et	par	de	en
Pain	le	que	et	quand	de	por
Mescreant	et	le	que	il	quand	si
Pecherresse	que	de	le	en	cel	quand
Merlin	le	que	et	quand	de	a
Florence	que	le	mais	a	et	quand
Anelés	que	le	et	mais	de	il
Buef	le	et	que	de	a	mais
Beguine	le	et	quand	que	a	de
Abesse	le	que	et	quand	il	seignor
Sauveur	le	quand	a	de	et	que

We should of course not expect the frequencies of line-initial lemmata to be completely consistent across each text of the J-corpus, since while an author may have preferences, as we have already acknowledged, they cannot be treated as a machine who will always give an identical output on the basis of a given set of inputs. That said, and on the understanding that an author's *œuvre* will never be truly homogeneous, it is possible to make some sense out of these apparently noisy data. By examining the corpus of texts as a whole, we may extract the lemmata which appear most commonly at the beginnings of lines across the entire J-corpus, and we can then compare the rates of use of each of these words for individual texts, which

will help us to identify whether there is indeed a unity to the frequency with which Jehan uses particular words at the beginning of each line.

Table 3.32 contains the total counts of the 6 words which appear most frequently within the J-corpus. These combined data appear to consolidate the trends observed in Table 3.31, namely that *le* is by far the most common lemma line-initially across the corpus, with *que* and *et* appearing next, before *de*, *quand*, and *a* form a group with reasonably similar rates.

Table 3.32 – the six most common lemmata at the beginning of a line across the entire J-corpus.

le	*que*	*et*	*de*	*quand*	*a*
825	505	482	292	283	279

By comparing the rates of the use of each of the above lemmata within each text with the rate at which the lemma is used in the rest of the J-corpus, we are able to calculate a p-value to determine whether there is a statistically significant difference between an individual text's use of a particular lemma in line-initial position, and the use observed across the rest of the J-corpus, given here in Table 3.33.[93] Lemmata with a p-value under the threshold of 0.05 display a use of a particular line-initial lemma that can be said with some certainty not to match that of the rest of the J-corpus, and are marked as achieving a PASS result.

Table 3.33 – results of the p-value tests comparing use of the six most frequent lemmata at the beginning of a line across the entire J-corpus.

Text	*le*	*que*	*et*	*de*	*quand*	*a*
Respon						
Chanoinnes						PASS
Pommes	PASS					
Romme	PASS			PASS		
Deux chevaliers						
Enfant rosti	PASS	PASS				
Povre chevalier			PASS			
Chevalier et escuier	PASS		PASS	PASS		
Narbonne						
Chevalier hermite						

[93] The full data for the tests can be found in Appendix 3B: Data from Table 3.48 to Table 3.53.

Cordouanier	PASS	PASS			
Juitel					PASS
Enfant sauva mere					
Eaue et vergier					
Pain		PASS			
Mescreant			PASS		
Pecherresse	PASS				
Merlin					
Florence	PASS	PASS	PASS	PASS	
Anelés	PASS	PASS			
Buef	PASS				
Beguine			PASS		PASS
Abesse	PASS				
Sauveur		PASS	PASS		PASS

The first observation of note from these tests is that the variability that we observed anecdotally in Table 3.31 appears to be borne out by the data. Only 7 texts do not display any statistically significant differences from the rest of the J-corpus for any of these lemmata, with 9 differing for a single lemma, 5 texts for 2 lemmata, and 2 texts (*Chevalier et escuier* and *Sauveur*) for 3 lemmata. *Florence* differs from the usage observed for the rest of the J-corpus in no fewer than 4 instances.

It is also the case that on occasion, even when a text makes use of a given lemma at a rate which locates it in an expected place in the order of the most frequent line-initial lemmata, its rate of use of that particular lemma still differs in a statistically significant way from that observed for the rest of the J-corpus. This is the case for *le*, for example, in *Pommes*, *Romme*, *Enfant rosti*, *Cordouanier*, *Buef*, and *Abesse*. In all except for *Buef*, where the rate is lower, this appears to result from the fact that *le* is used even more frequently in these texts than we might expect based on the rates observed across the J-corpus as a whole.

Taking a more global approach to the J-corpus, we are more successful in identifying possible ways in which Jehan's texts differ from trends observed elsewhere. Here we are able to

compare the corpus to four of our reference corpora, namely the P1-corpus of poems found in fr24432 which have not been attributed to Jehan, the F-corpus of *fabliaux*, the L-corpus of Marie de France's *Lais*, and the R-corpus of Rutebeuf's works. By comparing the J-corpus to all four corpora, we thus compare the J-corpus to a variety of corpus types, including works by a single author, works of a consistent genre, and works which are more disparate. In doing so, we observe the extent to which the texts of the J-corpus do, in fact, form a coherent corpus in how, collectively, they differ from other texts.

Table 3.34 – the six most common lemmata at the beginning of a line across the J-, P1-, F, L-, and R-corpora.

Corpus	1	2	3	4	5	6
J	*le*	*que*	*et*	*de*	*quand*	*a*
P1	*et*	*que*	*le*	*de*	*car*	*a*
F	*que*	*et*	*le*	*a*	*si*	*de*
L	*et*	*le*	*que*	*de*	*ne*	*a*
R	*que*	*et*	*le*	*ne*	*de*	*si*

Table 3.34 displays the six most frequent lemmata in line-initial position for the J-, P1-, F-, L-, and R-corpora. By contrasting our reference corpora with the J-corpus, we notice very much on a surface level that there is variation in which lemmata are the most frequent. While *et* is only the third most frequent lemma in line-initial position in the J-corpus, it is consistently the most frequent in the P1- and L-corpora, and appears in second position in the F- and R-corpora. Similarly, *quand* is the fifth most frequent line-initial lemma in the J-corpus, but drops from the list entirely for the other corpora, appearing in positions 13, 23, 12, and 11 for the reference corpora respectively, suggesting that Jehan's use of *quand* at the beginning of a line is unusually high, at least compared to the texts of the other corpora.

Considering the data from the other direction, we find a number of lemmata which appear with a high frequency in the reference corpora, but which do not appear as frequently in the J-corpus. *car* is the fifth most frequent lemma in line-initial position in the P1-corpus, but

appears in tenth position in the J-corpus. This lemma distinguishes the P1-corpus from the other corpora, however, rather than setting the J-corpus apart, since Jehan's use of this lemma is somewhat more typical, given that the F-, L-, and R-corpora contain *car* in positions 16, 20, and 10 respectively, and thus it is the use in the P1-corpus which is unusual.[94]

Similarly, *si* appears at a relatively high position for the F- and R-corpora, in positions 5 and 6 respectively, but only appears in positions 14, 10, and 16 for the J-, P1-, and L-corpora respectively. This could go some way towards suggesting a degree of similarity between the F- and R-corpora, which is unsurprising given the frequently similar content and form of many of the texts which they contain, but it also once again has the effect of separating the texts of the J-corpus from these reference corpora.

Simply ordering the most frequent line-initial lemmata in this way provides a useful yardstick by which to gain a sense of how our corpora may be distinct, but it ultimately does not allow us to reach any statistically assured conclusions. It is therefore also worthwhile comparing the J-corpus to our reference corpora in regards to the specific rates at which these lemmata are used at the beginning of a line.

Table 3.35 – rates of the use of the six most common lemmata at the start of lines in the J-corpus, compared to the P1-, F-, L-, and R corpora.

	J-corpus	P1-corpus		F-corpus		L-corpus		R-corpus	
Lemma	Rate	Rate	p-value	Rate	p-value	Rate	p-value	Rate	p-value
le	12.08%	5.60%	0	6.21%	0	7.99%	0	5.04%	0
que	7.39%	9.95%	0	12.05%	0	5.16%	0	13.81%	0
et	7.06%	10.04%	0	10.40%	0	8.66%	0.00086	9.26%	0
de	4.28%	4.71%	0.12936	3.00%	0.0004	4.87%	0.10957	3.84%	0.13383
quand	4.14%	1.84%	0	0.87%	0	1.95%	0	1.97%	0
a	4.09%	3.03%	0.00001	3.62%	0.20287	3.19%	0.00763	2.64%	0

[94] This high rate in the P1-corpus appears to be explained in part by the strong presence of this feature in a number of long allegorical texts within the corpus, namely *Dorée* (position 4), *Panthere* (3), *Arbre* (4), *Prise* (3), and *Miroir* (3), suggesting that this might be a feature common to this particular genre. This is not the sole explanation, however, since the lemma does also appear with relatively high frequency in other, shorter non-allegorical texts.

Table 3.35, containing the results of p-value tests for each of the six most common line-initial lemmata within the J-corpus, reveals that on the whole, line-initial lemmata are used significantly differently in the J-corpus from the rates observed in our reference corpora. It is unsurprising, given our observations above, that *quand* should have significantly distinct rates of use to the other corpora, but particularly notable is also the fact that there is a statistically significant difference for the three most frequent lemmata *le*, *que*, and *et*. While the relative positions of these terms were often comparable across the corpora, the rates at which these terms are used are not consistent. Thus we see that the J-corpus makes much greater use of *le* as a line-initial lemma than any of our reference corpora, while its use of *que* and *et* is lower than the rates observed for the reference corpora, with the exception of the unusually low rate observed for *que* in the L-corpus.

We are also able to draw broader conclusions from these data, namely that while uses of *le*, *que*, and *et* vary significantly across all the corpora we have examined, the rate of use of *de* appears to be more static, suggesting that this may be a feature which is determined by linguistic factors, rather than by the norms of a given genre, or the preferences of an individual author.

These comparative data therefore allow us to reach an interesting conclusion for the J-corpus: as we have seen, Jehan's use of line-initial words varies from text to text, with some poems displaying significantly different rates of use compared to the rest of the corpus. When taken together, however, the usage observed in the J-corpus does indeed appear to be significantly distinct from that found in texts by other authors, across a number of metrics. Thus, Munk Olsen (1963b: 152) is correct, of course, to argue that an author is not straightforwardly 'un cerveau électronique' with predictable outputs, but there do nevertheless appear to be macro-level trends which can unite a given corpus of texts through distinctiveness relative to some other corpus, which can contribute to the possibility of single-authorship.

Having established the validity of this method as a means of providing a sense of unity to the J-corpus, we may therefore also usefully examine the extent to which *Guillaume* and *Robert*, the two texts of the P1-corpus which we here also attribute to Jehan, line up with the trends observed for line-initial lemmata in the J-corpus. First, from Table 3.36, we observe that *Guillaume* in particular appears to line up relatively consistently with the J-corpus as a whole, with three of its six most frequently used line-initial lemmata matching those of the J-corpus. As for *Robert*, the text only lines up in two cases, yet the observed instances of *que* (113) and *le* (108) are sufficiently similar to suggest that we ought not to ascribe too much importance to the fact that these lemmata are swapped. A somewhat more significant observation in this text, however, is the presence of 'Robert' as the fifth most frequently use line-initial lemma, which contrasts with the texts found in the J-corpus in that none count proper nouns amongst their most frequent forms. *Robert* therefore shares some similar uses of line-initial lemmata with the J-corpus, but appears more peripheral than *Guillaume*.

Table 3.36 – the six most common lemmata at the beginning of a line in the J-corpus, Guillaume, *and* Robert.

	1	*2*	*3*	*4*	*5*	*6*
J-corpus	*le*	*que*	*et*	*de*	*quand*	*a*
Guillaume	*le*	*que*	*mais*	*et*	*quand*	*il*
Robert	*que*	*le*	*et*	*mais*	*Robert*	*a*

Looking again at the rates of use observed for the J-corpus and in *Guillaume* and *Robert*, we once again get the sense that for *Guillaume* in particular, the text is sufficiently similar to the broader J-corpus to support collective attribution. In Table 3.37 we observe that *Guillaume* displays rates of use with no statistically significant difference from the J-corpus for the lemmata *de*, *quand*, and *a*, three lemmata which were shown to have statistically different rates when the J-corpus was compared to our reference corpora, in line with a number of texts of the J-corpus as displayed in Table 3.33. While there are differences in usage, then, in the case

of *Guillaume* it appears likely that these are the result of variation within a corpus which is nevertheless itself distinct from other texts.

Table 3.37 – rates of the use of the six most common lemmata at the start of lines in the J-corpus, compared to Guillaume *and* Robert.

	J-corpus	*Guillaume*		*Robert*	
Lemma	Rate	Rate	p-value	Rate	p-value
le	12.08%	15.08%	0.00831	10.63%	0.18122
que	7.39%	9.60%	0.01713	11.12%	0.00009
et	7.06%	4.54%	0.00381	5.02%	0.01699
de	4.28%	3.16%	0.10692	2.95%	0.046
quand	4.14%	3.59%	0.41692	2.66%	0.02275
a	4.09%	3.06%	0.12812	3.84%	0.71074

Robert, however, again appears somewhat more problematic. The text displays statistically significant differences from the J-corpus for four lemmata, namely *que*, *et*, *de*, and *quand*. If we compare this test to those performed on the individual texts of the J-corpus, we therefore find that *Robert* sits alongside *Florence*, which is also distinct from the rest of the J-corpus in the case of four lemmata, and indeed overlaps with it for three of these.

One might therefore contemplate the exclusion of both *Florence* and *Robert* from Jehan's *œuvre*, but it is once again worth reiterating the fact that an author's output is unlikely ever to be wholly homogeneous and predictable. Where on the one hand there may be evidence of difference within a large selection of an author's *œuvre*, similarities may be seen in sub-groupings of this collection. One final example will allow us to suggest that this is true for Jehan's texts in the case of line-initial lemmata, and to gain a surer footing from which to attribute *Robert* to Jehan, namely the fact that the lemma *mais* only appears in the top 6 lemmata for five texts which we are more or less confidently attributing to Jehan, namely *Florence*, *Anelés*, and *Buef* from the J-corpus, as well as *Guillaume* and *Robert*. All of these texts also distinguish themselves from the rest of Jehan's texts through the fact that they are far longer than any other texts of the J-corpus, forming what we may term the 'long-text'

sub-group. That *mais* is used a great deal less in the shorter texts of the J-corpus suggests that Jehan's word choices may have varied depending on features of the text he was writing, such as its intended length. This, along with the fact that *Robert* lines up with *Florence* in its use of line-initial lemmata, suggests that what matters less than *Robert*'s points of difference from the J-corpus as a whole is therefore the fact that the text shares similarities with other texts from the 'long-text' group, allowing us to strengthen our attribution of both *Guillaume* and *Robert* to Jehan. We shall observe further evidence for the validity of this subdivision in the next section.

Not only has the method pursued in this section therefore allowed us to affirm the collective attribution of these texts, it has also provided further evidence for the claim that an author cannot be simply treated as a machine who, given a certain input, will always produce the same output. It is therefore worthwhile considering other types of shared structures across the J-corpus which display both the diversity and unity of Jehan's texts.

3.7.3 Onset of speech

A further object of analysis proposed by Munk Olsen (1963b: 150) for the examination of common structures, but once again not widely applied by him across the J-corpus, is the moment of the onset of discourse in the line.[95] He proposes a typology for the ways in which direct speech is introduced within Jehan's poems, identifying three types, as below:

> Type A: Discourse begins at the start of a new line, with the speaker indicated in the previous line.

[95] Here we talk of the 'onset of discourse' or 'discourse onsets' to refer to that moment at which direct speech begins within a text, as marked in modern editions with an opening set of quotation marks. A great deal of useful work has been done on different kinds of discourse, the analysis of which falls outside our scope here (see, for example, Ducrot (1984, 1989), Cerquiglini (1981, 1989) and Marnette (1998, 2005)). Avenues for further examination within the J-corpus in particular include the frequency of the use of direct discourse compared to other kinds of discourse, details of the identity of the speaker, and where in a text discourse appears.

Type B: Discourse begins at the start of a new line, with the speaker indicated later in the same line, before discourse resumes.

Type C: The speaker is indicated at the beginning of a new line, followed by the onset of discourse.

Examples of each of these types abound:

Type A:
Le chevalier l'apelle et si l'araisonna :
"Amis, celi t'en gart qui le monde forma !" (*Chevalier et escuier*, vv. 25–26)

Type B:
"Sire," dit la pucelle, "je ne vous desdi mie." (*Respon*, v. 102)

Type C:
Le chevalier respont : "Voulentiers le ferai." (*Chanoinnes*, v. 136)

There is, however, little concrete evidence provided by Munk Olsen to support his claims as to the frequency of use of these types, and it is equally unclear to what extent the corpus is unified in the texts' relative use of them. A corpus-wide approach, aided by computational and statistical methods, can therefore shed further light on the matter.

The first observation from an examination of all onsets of discourse across the J-corpus is that Munk Olsen's types are too restrictive, and do not encapsulate all instances. For Type A, for instance, we notice that when direct discourse begins at the start of a new line, the speaker is not necessarily identified in the preceding line, and may in fact have been identified earlier, as in a passage from *Pain*:

Li ennemis parole a l'ange d'une part
Por quoi Diex veult avoir en cestë ame part :
"Il m'est avis qu'il s'est aperceüs trop tart, […]" (*Pain*, vv. 113–115)

For Type B, in cases where the direct discourse given by a character is particularly short, direct discourse may indeed begin at the start of a new line, with the speaker's identity given afterwards, but discourse does not then resume:

> "Non ferai je," dist cil, qui bien la sot entendre. (*Beguine*, v. 46)

It is also possible, despite Munk Olsen's suggestions to the contrary, for direct discourse to resume in a following line, rather than within the same line as the initial onset:

> "Par ma loy," dit li peres, qui fu fel et engrez,
> "Qu'il a passé .vij. ans, nous le suivrons de prez.
> […]" (*Juitel*, vv. 37–38)

As for Type C, it is not always the case that when the introduction of a speaker occurs before but in the same line as their direct discourse, this introduction is the first element of the line, since there are examples of elements of narration appearing as the first element of such lines:

> L'ostel vit d'un hermite, si a dit : "Diex aïde !" (*Sauveur*, v. 180)

Indeed, on occasion, the first element of a line may be the end of a previous unit of discourse:

> "[…]
> Qu'il m'ont fait." La voiz dist : "Tu ne fais pas folie." (*Florence*, v. 724)

Beyond the types proposed by Munk Olsen, we also observe a further means of introduction of direct discourse, namely the often stichomythic juxtaposition of two units of speech with no intervening element, with the switch in speaker usually identifiable in the written form only through the content of the discourse.[96] These onsets of discourse may take place between one line and the next, or within the same line:

[96] We could take this type of discourse onset as evidence for the texts' possible orality, since it would be far easier to identify a change in speaker through the actions and tone of voice of a storyteller than it is on the manuscript page, with fr24432 only rarely containing any punctuation to indicate the beginning of a unit of discourse. On

"Dame, vous estiez morte. Qui vous resuscita ?"
"Eu non Dieu, dous amis, cele qui Dieu porta." (*Cordouanier*, vv. 131–132)

"Sire, quel non avez ?" "On m'appelle Merlin." (*Merlin*, v. 54)

Finally, examples can be found in which more than one type of discourse onset intersects:[97]

[…] Et celle respondi :
"Certes, sire," dist elle, "ne vous mentiré mie.
[…]" (*Beguine*, vv. 68–69)

We may therefore amend Munk Olsen's typology to the following:[98]

Type A: Discourse begins at the start of a new line, with the speaker indicated in an earlier line.

Type B: Discourse begins at the start of a new line, followed by an indication of the speaker later in the same line. The discourse may then resume, either later in the same line, or in a following line.

Type C: The introduction of the speaker is followed in the same line by the onset of discourse.

Type D: Discourse begins immediately following the end of a previous unit of discourse by another speaker, either as a continuation of the same line or at the beginning a new line.

With our amended typology, we can thus examine the extent to which the use of each type is consistent across the J-corpus. From the data in Table 3.38, we observe that of a total of 739 onsets of discourse, Types B and C count for roughly a third each, while the totals for Types A and D are much lower. Munk Olsen's claim that Type B is 'de beaucoup le plus fréquent'

these rare occasions where such a feature is marked, it is usually with a *punctus*. For more on the use of punctuation to introduce direct discourse, see Marnette (2006).

[97] When this occurs, both types of discourse onset are included in our counts.

[98] Types A and C combine to form a category of discourse onset which corresponds with what Cerquiglini (1981) and Marnette (2006) have called *prolepse*, that is, the introduction of a speaker to direct discourse before it begins, while Type B is *analepse*, being the introduction of a speaker after direct discourse has already begun.

(1963b: 160), in a simply numerical sense at least, therefore appears to be overblown, since while onsets of Type B are indeed the most frequent kind, they cannot be said to be far and away the preferred method, given how frequently we also find onsets of Type C.

Table 3.38 – counts of each type of discourse onset for each text in Munk Olsen's (1978) corpus of Jehan's dits.

Text	Type A	Type B	Type C	Type D	Total
Respon	6	11	10	6	33
Chanoinnes	11	8	12	1	32
Pommes	11	19	9	7	46
Romme	6	10	1	0	17
Deux chevaliers	5	16	7	1	29
Enfant rosti	3	11	4	0	18
Povre chevalier	4	1	7	12	24
Chevalier et escuier	4	19	11	2	36
Narbonne	5	16	3	0	24
Chevalier hermite	6	9	9	1	25
Cordouanier	3	10	5	2	20
Juitel	9	2	5	4	20
Enfant sauva mere	4	9	13	0	26
Eaue et vergier	3	4	2	0	9
Pain	4	5	7	0	16
Mescreant	0	11	12	0	23
Pecherresse	2	2	6	0	10
Merlin	3	3	12	20	38
Florence	13	9	50	0	72
Anelés	11	3	44	1	59
Buef	13	23	17	1	54
Beguine	5	11	5	0	21
Abesse	3	23	4	3	33
Sauveur	15	32	5	2	54
	149	267	260	63	739

By considering the relative rates of use of each type per text, as represented in Figure 3.61, it becomes clear that there is a great deal of variation in the types used by individual texts, in particular when considering Type D.[99] With the exception of *Mescreant*, which has no examples of Type A, all texts contain examples of Types A, B, and C. For Type D, however,

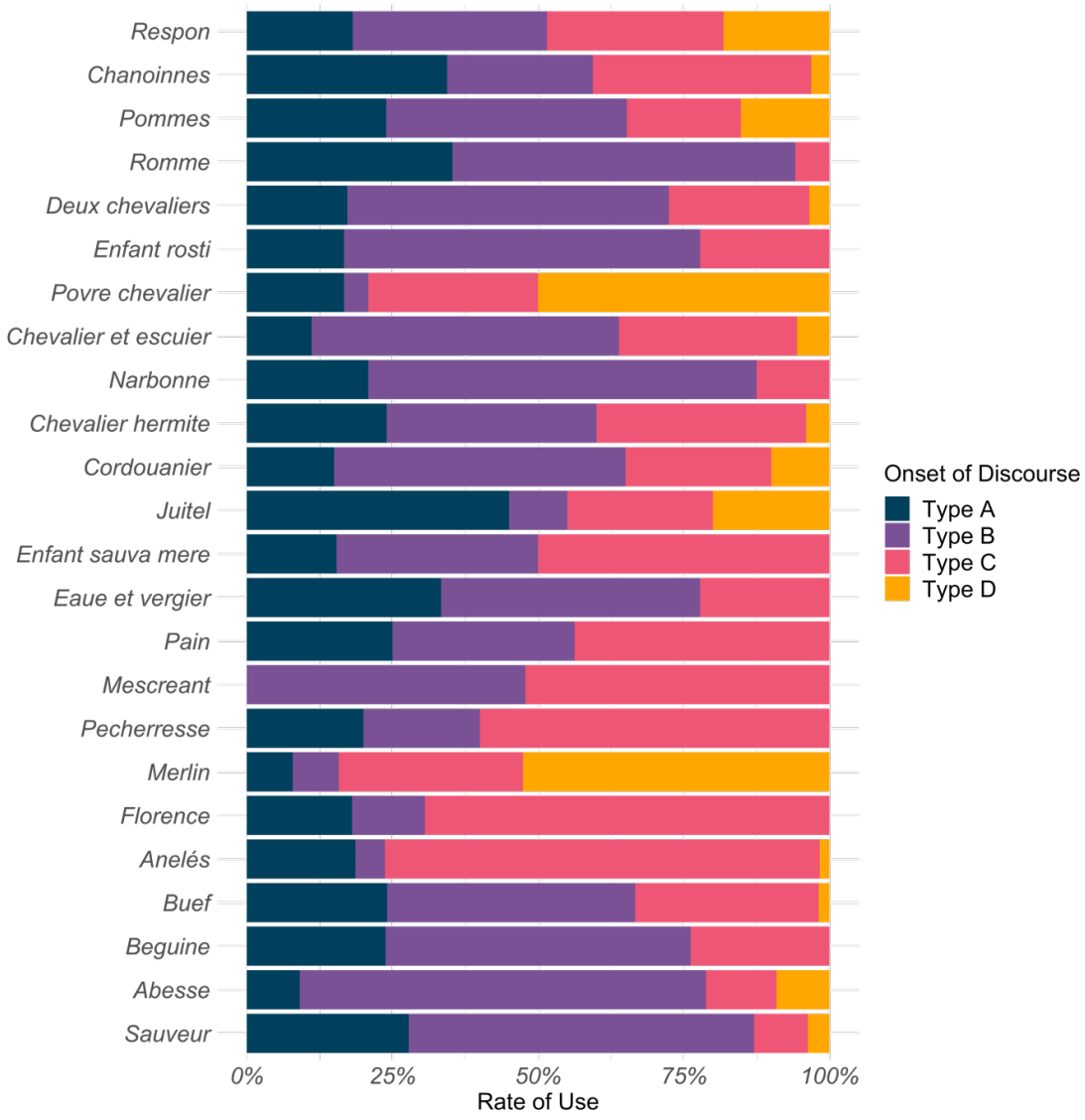
[99] The full data can be found in Appendix 3B: Data from Table 3.54 to Table 3.57.

10 of the 24 texts contain no examples of this type at all. In *Povre chevalier* (50%) and *Merlin* (52.6%), by contrast, at least half of the total onsets of discourse can be classified under Type D. There does not appear to be a narrative explanation for this distinction, either, since while *Povre chevalier* and *Merlin* are indeed texts which contain a lot of the fast-paced dialogue that might lend itself to unindicated switches in speaker, so too are many other texts of the corpus, such as *Florence*, which contains no such examples.

Of further note is that *Florence* and *Anelés* in particular appear to favour Type C onsets, almost to the exclusion of all other types, resulting in a much lower than expected use of Type B. Given that these texts are two of the longest in the corpus, and contain the greatest number of discourse onsets, we are able to reinforce our hypothesis above that the longer texts within the J-corpus are more likely to differ from the rest of the corpus than others. Indeed, if one excludes both *Florence* and *Anelés* from the data, then Munk Olsen's suggestion that Type B is 'de beaucoup le plus fréquent' (1963b: 160) appears far more plausible, since we observe 255 instances of Type B, but only 166 instances of Type C. Notably, however, *Buef*, the other text within the J-corpus belonging to the 'long-text' sub-group, displays rates of use of Types B and C more in keeping with the rest of the corpus, suggesting that the uses observed in *Florence* and *Anelés* are not an established feature of the whole of this sub-group.

Figure 3.61 presents these data as rates, and reveals that there is a great deal of variability in the use of discourse onsets within the J-corpus, very little of which can be satisfactorily explained with reference to the specific narrative requirements of a given text.

Figure 3.61 – rates of use of each type of discourse onset in the J-corpus.



Furthering this examination through the use of p-value tests, we are able to confirm our observation that while the texts are relatively comparable in their use of certain types of discourse onset, they differ quite dramatically in their use of others. In this analysis, a `PASS` result to the p-value test suggests that there is at least a 95% chance that the way in which a particular discourse onset type is used in one individual text is statistically distinct from that of the remainder of the corpus. Looking at the results of the tests in Table 3.39, we observe that the use of quotations of Type A appears reasonably consistent across the corpus, with only 3 texts displaying a statistically significant difference. The same cannot be said, however, for

Types B and C, with texts differing from the expected rates of the rest of the corpus on 11 and 7 occasions respectively. Of particular interest is the run of 5 texts (*Deux chevaliers* to *Narbonne*), which appear contiguously in fr24432, and which are all distinct from the norm seen elsewhere in the corpus. We should not be tempted to claim that this is because they all use discourse onsets of Type B in the same way, however, since we can observe from Figure 3.61 that while *Deux chevaliers*, *Enfant rosti*, *Chevalier et escuier*, and *Narbonne* display what appears to be an unusually high rate of usage of Type B, *Povre chevalier*, the third text in the sequence, displays the opposite, using discourse onsets of Type B surprisingly infrequently.

Table 3.39 – results of the p-value tests comparing use of the each type of discourse onset across the J-corpus.

Text	Type A	Type B	Type C	Type D
Respon				PASS
Chanoinnes	PASS			
Pommes			PASS	
Romme			PASS	
Deux chevaliers		PASS		
Enfant rosti		PASS		
Povre chevalier		PASS		PASS
Chevalier et escuier		PASS		
Narbonne		PASS	PASS	
Chevalier hermite				
Cordouanier				
Juitel	PASS	PASS		
Enfant sauva mere				
Eaue et vergier				
Pain				
Mescreant	PASS			
Pecherresse				
Merlin		PASS		PASS
Florence		PASS	PASS	PASS
Anelés		PASS	PASS	PASS
Buef				
Beguine				
Abesse		PASS	PASS	
Sauveur		PASS	PASS	

With regards to Type C, of note is the fact that, with the exception of *Florence* and *Anelés*, which as we have already seen have an unusually high rate of usage of this type, all 5 other examples pass the p-value test as a result of significantly lower than average usage. We may also note that a statistically significant difference in the use of Type C also often appears to correlate with a difference in the use of Type B, with 5 texts overlapping in this way (*Narbonne*, *Florence*, *Anelés*, *Abesse*, and *Sauveur*), in all cases because a high rate of use of one type is counterbalanced by a low rate of use of another. This could perhaps suggest that Types B and C were somehow seen as variants in alternation, with a high rate of use of one type leading to fewer possibilities for the other, whereas the use of Types A and D was more likely to be conditioned by other factors, such as narrative or versificatory requirements. While this would be a surprise, given the cognitive difference between the proleptic Type C and the analeptic Type B,[100] this can perhaps be explained by the fact that, particularly in the alexandrine lines of the J-corpus, discourse onsets of Types B and C are usually those which take up roughly half of an individual hemistich. They are thus roughly interchangeable from a versificatory perspective, and so were an author to decide to make frequent use of one type for a given text, then the opportunities for the other would be significantly diminished.

Turning finally to Type D, we observe that in addition to the unusually high rates of use in *Povre chevalier* and *Merlin*, the rate of use in *Respon* is sufficiently elevated to result in a `PASS` result, suggesting some unusual use of Type D onsets in this text when compared to the whole corpus. *Florence* and *Anelés* are again united in the fact that, despite their long length and numerous onsets of direct discourse, there is not a single Type D onset in the former, and only one in the latter.

[100] See note 98.

Given these points of similarity between *Florence* and *Anelés*, we are reminded of Morawski's claim that 'nous croyons que [*Florence* and *Anelés*] sont l'œuvre d'un même auteur, mais nous n'avons pas acquis la conviction qu'ils soient de Jean' (1939: 351–352). I disagree with Morawski's claim here on the basis that the data here do not go so far as to encourage the exclusion of these texts from the J-corpus, given the unifying features presented elsewhere in our analysis, but our analysis has nevertheless enabled us to encounter further stylistic similarities between these two texts, suggesting that if they are the work of the same author, then they may have originated from a similar context of composition.

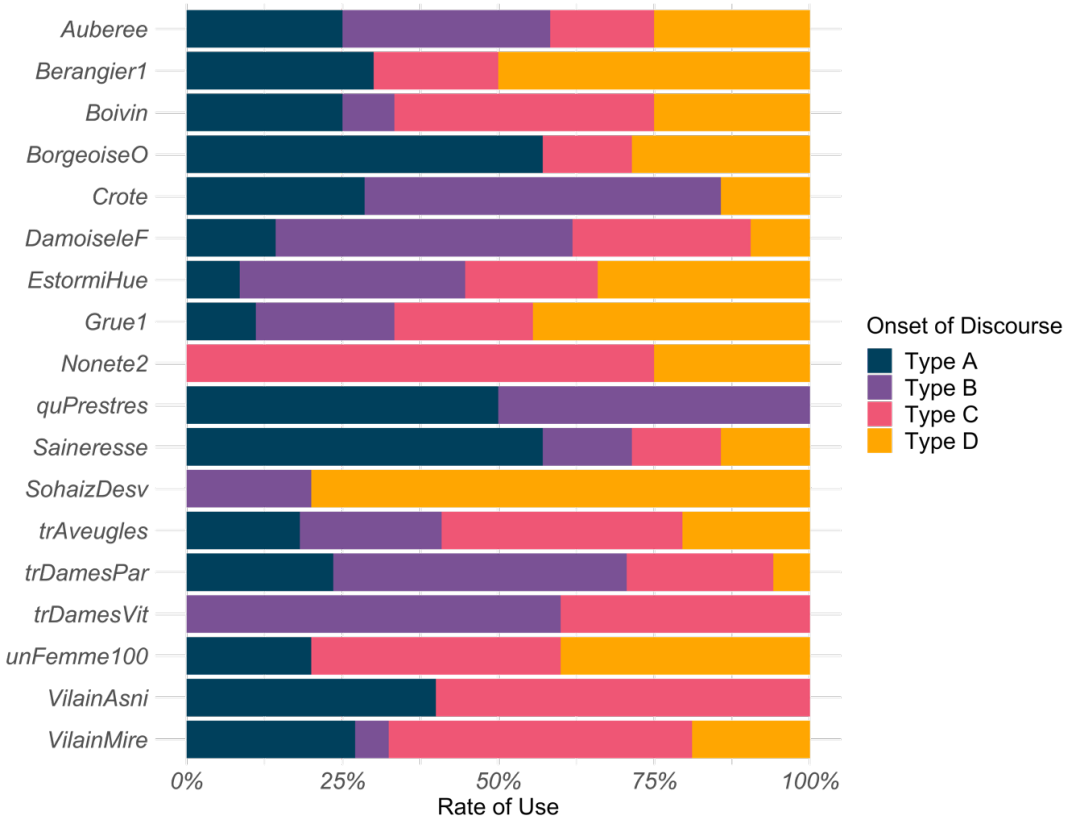
As with our examination of line-initial lemmata, what is most noticeable about these granular results is that the texts of the J-corpus often vary significantly in their use of a particular type of discourse onset. While this makes it difficult to use these data alone to strengthen the hypothesis of shared authorship across the J-corpus as a whole, it does nevertheless create interesting connections between texts, such as *Povre chevalier* and *Merlin*, which we would otherwise be unlikely to link together.

Again, when we compare the J-corpus with our reference corpora, it becomes immediately apparent that there is more cohesion within the J-corpus than the text-level data would suggest. Figure 3.62 displays the combined rates of use of different types of discourse onset for the texts of the F-corpus, with the exception of four texts which contain either 1 or 2 instances of direct discourse.[101] The most obvious observation to be made on the basis of these data, when compared to those of the J-corpus, is that the texts of the F-corpus make much greater use of Type D onsets, to the point that this distinction is strongly statistically significant

[101] For brevity in the figure, I have trimmed the long names given to some of these *fabliaux* by the texts' encoders (Pierreville 2023). For full titles and details of the manuscripts from which the encodings are drawn, see Appendix 3A: F-corpus contents. It is worth noting that for the F-corpus, Type B discourse onsets also include those where discourse begins on one line, and the speaker is introduced in the following line. This does not appear in the J-corpus, but this is unlikely to have been the result of authorial choice, rather it appears to be a function of the octosyllabic metre of the *fabliaux*, since there is less space than in the alexandrines of the J-corpus in which to fit both the direct discourse and its analeptic discourse onset.

between the corpora as a whole (p = 0), and the same can be said for Type B onsets between the two corpora (p = 0.00514).

Figure 3.62 – rates of use of each type of discourse onset in the texts of the F-corpus containing 5 or more instances of direct discourse.



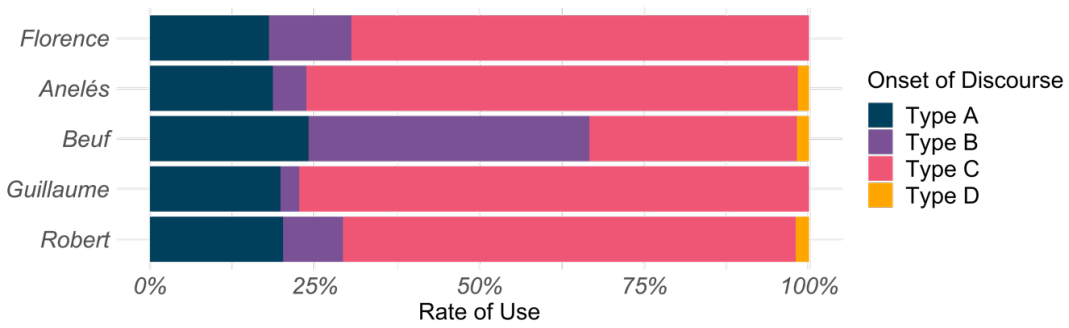
Unlike in the J-corpus, meanwhile, where there was only one example of a text which did not contain at least one discourse onset from each of Types A, B, and C, the presence and absence of such forms is a lot more variable in the F-corpus, with 3 texts not containing Type A onsets, 5 omitting Type B onsets, and 2 omitting Type C onsets. On two occasions, in the cases of *Nonete* (*Nonete2*) and *Li Sohait desvez* (*SohaizDesvezBodel*), the texts only contain two types of discourse onset.

Compared to the J-corpus, then, the F-corpus of texts is far more variable. Given that the latter texts are consistent in genre but do not share an author, we are able to conclude that the use of discourse onsets, while somewhat influenced by genre and form, appears also to be

determined by authorial preference. We have thus once more been able to confirm for the J-corpus that what looks at first like significant variation within the corpus, when considered in the context of a reference corpus, instead ultimately allows us to strengthen the sense of unity of the corpus. That is not to say, of course, that the texts of the J-corpus are not significantly diverse in their approach to the onset of discourse. It is simply the case that for most of the texts it contains, and certainly in a holistic sense across the corpus as a whole, this variation is not as significant as that which we may expect to find elsewhere in corpora of texts which do not share a single author.

Turning briefly to *Guillaume* and *Robert*, and comparing the rates within them to the three texts of the 'long-text' grouping which we include within the J-corpus, we observe from Figure 3.63 that the two texts outside the J-corpus are almost exactly comparable to *Florence* and *Anelés* in their use of discourse onsets. All five make very similar use of Type A onsets, and *Florence* and *Guillaume* contain no examples of Type D onsets, while *Anelés*, *Buef*, and *Robert* contain only very few. As a result, the outlier within the grouping is not, in fact, either of the texts which appear outside of the J-corpus, rather it is *Buef* which distinguishes itself from the others through its increased use of Type B onsets, and decreased use of Type C, whereas the other texts are consistent in their significant preference for Type C over Type B.

Figure 3.63 – rates of use of each type of discourse onset in the texts of the 'long-text' sub-group of the J-corpus, alongside Guillaume *and* Robert.



Thus, while *Guillaume* and *Robert* are distinct from the J-corpus as a whole for three out of the four discourse types,[102] just as we found with *Florence* and *Anelés*, the data are nevertheless reassuring in their suggestion that the former two poems are entirely coherent with the bulk of the 'long-text' sub-group within the J-corpus, allowing us to suggest with further confidence that these texts are also the work of Jehan.

As a final point on the range of similarities in structure which we have observed for the J-corpus and beyond, it is worth once more revisiting Munk Olsen's observation (1963b: 152) that an author ought not solely to be thought of as 'un cerveau électronique dans lequel une impulsion de contenu donnée déclenche automatiquement et infailliblement l'impulsion d'expression correspondante.' As we have seen for both line-initial words and the onset of discourse, if any conclusions are to be drawn at all from such data about whether a group of texts share a single author, then this must be done alongside an acknowledgement that the mind of an author is never fixed; no text is ever written with the same combination of inspiration and creativity as any other, and so it is far from surprising that variation should be inevitable, even in the works of a single author.

On the whole, however, we have observed that there are a great many similarities of vocabulary and structure across the texts of the J-corpus, both those determined by Morawski, Kleist, and Munk Olsen, and those uncovered using statistical and digital methods which were unavailable to them. While each of these similarities are not in themselves sufficient to prove shared authorship, they do once again contribute to the unity of the J-corpus alongside *Guillaume* and *Robert*, allowing us to be more assured in their attribution to Jehan.

[102] *Guillaume*: A: $p = 0.93046$; B: $p = 0$; C: $p = 0$; D: $p = 0.00234$. *Robert*: A: $p = 0.99286$; B: $p = 0$; C: $p = 0$; D: $p = 0.02269$.

3.8 Language

3.8.1 Localisation

Much of the possible linguistic analysis of the J-corpus has already been carried out, most recently and in most detail by Munk Olsen (1963a, 1978: cxii–cxlii). These analyses have largely centred around questions of Jehan's apparent dialect, and the extent to which this can be used as a means of localising the texts. Certainly, given Jehan's byname, as apparently originating from Saint-Quentin, we might be tempted to entertain the possibility that his own dialect was one closer to Picard. Fundamentally, however, as Munk Olsen (1963a: 124) acknowledges, 'à partir [des] traits de langue attestés comme authentiques par la rime ou par la scansion, il est bien difficile, voire impossible, de localiser de façon précise l'auteur des *Dits*.' At best, it seems, Jehan's language can be best described as containing a mixture of traits associated with a number of different, predominantly northern and western, dialects.

Morawski (1939: 333) categorises the language of the J-corpus as being primarily a 'mélange de formes franciennes avec des formes septentrionales et occidentales,' which he notes is a common feature of many of the *dits* from this period written in alexandrine rhyming quatrains. In support of this, Morawski's evidence is primarily morphosyntactic, noting in particular the frequent presence of the 1.PL ending *-on*, common in western and Picard dialects, as well as picardisms such as the presence of syllabic, epenthetic *e* within the future stem, as in *penderai* (*Cordouanier*, v. 106) and the interchangeable use of the 2.PL possessives *vo* and *vostre* (1939: 333).[103] Munk Olsen furthers the observation of Picard traits in the J-corpus, in particular the fact that the tonic and atonic *e* pair in *-iee* is often reduced to a single tonic vowel *-ie*, as well as noticing an isolated example of the Picard form *fiex* 'son' (*Anelés*, v. 637), which is unlikely to be a scribal modification given its rhyme with the forms *iex*, *viex*, and *ciex*.

[103] We ought to note, though, as Munk Olsen (1963a: 124) does, that this latter feature was also widespread outside of Picardy towards the beginning of the fourteenth century.

That said, as Munk Olsen (1963a: 124) observes, these picardisms in the J-corpus appear to be very much the exception to the rule, since 'la plupart des particularités du texte sont caractéristiques de l'Ouest.' These include the frequent reduction of diphthongal /ie/ to /e/ (as in the presence of quadrisyllabic *appareillierent* rhyming with forms without /i/ such as *porterent*; *Buef*, vv. 670–671), monophthongisation of pre-consonantal /ai/ to /ɛ/ (as in *nestre* rhyming with *prestre*; *Narbonne*, vv. 144–145), and the raising of tonic /o/ to /u/ (as in the rhyme of *seignor* with *dolour*; *Chevalier et escuier*, vv. 217–218), alongside the use of forms such as *el*, *li*, and *celi*. It is of course possible that many of these forms could result from scribal changes, but their preponderance in instances where such a change would have resulted in inconsistencies in the rhyme or metre allows us to say with some degree of certainty that this is not always the case, and that the language of the J-corpus thus exists somewhere between Picard and western varieties.[104]

Munk Olsen (1963a: 125) hence suggests that:

> La meilleure solution à ce problème inextricable serait certainement d'y voir une langue artificielle et hybride : l'auteur a pris des notes dans n'importe quel dialecte […], et, puisqu'il n'avait aucun sens d'une langue normalisée, il n'avait aucun scrupule à les utiliser même s'ils étaient en contradiction avec son propre dialecte qui a pu être le picard aussi bien qu'un des ceux du centre occidental de la France.

There are a great many questionable assumptions in this claim, not least that Jehan would have had no concept of a normalised, or prestige, dialect. That said, we may interpret Munk Olsen's underlying claim to be the more neutral suggestion that there is evidence of diglossia in the texts of the J-corpus, with the author perhaps usually speaking a Picard dialect which influences his predominantly western *scripta*. Whether or not this is true, through these observations we reach a more valuable conclusion for our purposes here, namely that the texts of the J-corpus present a degree of unity in the extent to which they combine Picard,

[104] This observation also applies to *Guillaume* and *Robert*, which display a great many of these features.

western, and *francien* forms, regardless of the reasons why such a combination exists, which at the very least does not discourage collective attribution.

3.8.2 Dating

As for the dates of the texts in the J-corpus, there is little concrete evidence to come from the language itself. However, Morawski (1939: 334) suggests that the poems were most likely written within the period 1313–1345, the range of dates given within the texts and paratexts of fr24432 for texts outside of the J-corpus, offering as evidence the fact that analogical forms such as non-epicene *grande* do occasionally appear in rhyming position or at a point in the line where the two-syllable form is metrically necessary (once each in *Narbonne*, *Chevalier et escuier*, and *Buef*, as well as in *Guillaume* and *Robert*), but are still rare compared to the non-analogical forms.

Munk Olsen, meanwhile, suggests an earlier date range than Morawski, but nevertheless acknowledges the difficulty of determining any precise date of composition, suggesting that 'nous ne sommes guères plus avancés quoique nous n'ayons que le choix entre la fin du XIII^{e} et le début du XIV^{e} siècle' (1963a: 125). Evidence of forms associated with the late thirteenth century comes in the form of enclisis in *nel*, *nes*, and *sel*, the persistence of final *-t* in *-ent*, and the occasional use of disyllabic imperfect and conditional endings (1978: cxxxv–cxxxvi), while his arguments for placing the texts into the fourteenth century come again from the appearance of analogical developments such as *grande*.

In addition, Munk Olsen observes what he argues is a low level of degradation of the case system, particularly for an author with a predominantly western *scripta*, with a reasonably consistent rate of around 10% of nouns displaying what might be considered the 'incorrect' case, for those nouns where case can be definitively said to have been determined by the author, rather than a scribal (that is, where nominative and oblique forms are imparisyllabic).

Ultimately, then, while there are general trends to be observed within Jehan's language which increase the sense of unity of his corpus, these do not offer a great deal of clarity as to exactly where, or exactly when, he was writing. In order to continue our analysis of the shared-author hypothesis, we therefore turn to elements of the language of the J-corpus which do potentially have more to offer such questions.

3.8.3 Possessive adjectives

For Morawski, one of the most important linguistic features of the texts of the J-corpus that he examines, which in his eyes can be used to unify the corpus and exclude possible external contenders, is the fact that, in his smaller corpus of Jehan's *dits*, there are no instances of *mon*, *ton*, or *son* used before a feminine singular noun beginning in a vowel (Morawski 1939: 334, 1940: 506, 509, 516, 518). This is despite the fact, he argues, that this was increasingly possible in a number of dialects at the time that these texts were written, with Jehan preferring instead to use just the elided forms *m'*, *t'*, and *s'*. Munk Olsen (1963b: 155), while critical of Morawski's reliance on this single feature as a means of attributing authorship to Jehan, does not disagree with his analysis, arguing that, by the fourteenth century, use of *mon*, etc. before a feminine vowel-initial noun was 'tout à fait courant,' and that this is therefore an unusual feature within the J-corpus.

While surprisingly little linguistic scholarship exists on this feature, it would appear that the absolutist nature of these claims does not necessarily fit with the observed trends. Lefèvre *et al.* (1976–1983, ii: §220) note only that, as Old French moves towards Middle French, 'les féminins *ma* et *sa* s'élident encore [devant voyelle] malgré un recours de plus en plus fréquent aux masculins *mon*, *son*,' which could be seen to fit broadly with Morawski and Munk Olsen's claims. Marchello-Nizia *et al.* (2020: 706, 710) offer a little more detail, suggesting that over the course of the fourteenth century 'les formes atones élidées sont remplacées analogiquement par les formes du masculin,' but that this process of evolution depends on the

noun which the possessive is describing. For example, the elided form of *ma* etc. is common before *ame*, *amie*, *amour*, and *espouse* even towards the end of the fifteenth century, while *mon* etc. predominate for other nouns as early as the middle of the fourteenth century, becoming the norm by the beginning of the fifteenth.

Thus, if we are to take the texts of the J-corpus as having been written either at the end of the thirteenth century, or early into the fourteenth century, then arguably we should not be surprised to see no forms of *mon*, etc. preceding feminine vowel-initial nouns, even those for which the possessive system developed over a shorter period of time. This is especially the case given Marchello-Nizia's suggestion (1979: 138) that this change took place more slowly in Picard dialects, such that we may expect an even later appearance of this feature in texts like those of the J-corpus which contain Picard features. As ever, digital textual analysis has a role to play here, enabling us to take a more global view of the J-corpus, and to place it more clearly within both its manuscript and literary contexts, nuancing both Morawski and Munk Olsen's claims, as well as those of the linguistic scholarship.

Of course, finding an instance of *mon* etc. before a feminine vowel-initial noun would not be enough to claim that a text was not written by Jehan. The preferences of an individual author will inevitably change over time, and the possible alternation of *mon* etc. with *ma* etc. for reasons of metre means that it is entirely possible that such decisions were not solely due to authorial preference.

That said, on the whole, Morawski's observation regarding Jehan's preference for elided *ma*, etc. does appear to be borne out by a holistic analysis the corpus. In general, *ma* etc. appears before feminine vowel-initial nouns, and *mon* etc. is restricted to use before masculine nouns. Examples from the J-corpus of each use can be seen in Table 3.40 and Table 3.41.

Table 3.40 – examples of elided ma *etc. preceding feminine vowel-initial nouns within the J-corpus.*

Text	Line	LH	KW	RH
Deux chevaliers	106	A toi	m'	ame et mon corps commant et a Jhesu
Buef	617	"Par	t'	ame pren y garde," et celle va tantost
Mescreant	12	Jamais	s'	ame n'avra vers Dieu redemption

Table 3.41 – examples of mon *etc. preceding masculine vowel-initial nouns within the J-corpus.*

Text	Line	LH	KW	RH
Pecherresse	20	Nient plus qu'a	mon	enfant, ne li vouldrai faillir
Narbonne	196		Ton	enfant te vois querre en santé et en vie
Pommes	328	Le seigneur se restorne,	son	enfant aperçoit

In most cases of the use of *mon* etc., the masculine gender of the noun it precedes can be clearly established, either by convention or through the presence of surrounding words with masculine agreement.[105] The only case for which this is not true is a single instance of 'mon errour' found in *Chevalier hermite* (v. 48). The 2020 *Dictionnaire du Moyen Français* (DMF) lists this noun as variable between masculine and feminine, but it is not possible to determine the gender of this word in the J-corpus by examining the contexts in which it appears. In all 3 uses of *erreur* across the J-corpus, given in Table 3.42, we can see that there are no surrounding morphosyntactic clues that could provide this information. Even the use of 'grant' is not definitive since, as we have seen, within the J-corpus this form appears to be in alternation with the analogical form *grande* which later comes to replace it.

Table 3.42 – uses of erreur *within the J-corpus.*

Text	Line	LH	KW
Pommes	162	Du preudomme dirai qui fu en grant	errour
Pommes	265	Li varlés retorna qui fu en grant	errour
Chevalier hermite	48	Veil aler a confesse regehir mon	errour

[105] The term *afaire*, for example, which appears with *mon* etc. in both *Florence* (v. 237) and *Sauveur* (v. 201), is listed in the DMF as alternating between masculine and feminine, but it can be straightforwardly determined to be masculine in these examples, being accompanied by forms such as *le*, *bon* and *cest* on a number of occasions elsewhere in the corpus, including within *Sauveur* (*cest*: v. 5, *le*: v. 10).

We are therefore forced to conclude at the very least that the gender of *erreur* in these cases is unknowable. While it is therefore possible that the example in *Chevalier hermite* is indeed a use of *mon* etc. before a feminine vowel-initial noun, the ambiguity of this example means that it would be insufficient on its own to cause us to reject Morawski's claims about Jehan's possessive usage. *Guillaume* and *Robert* also contain no examples of *mon* etc. before a feminine vowel-initial noun, and as such we may conclude that across all the texts we have thus far attributed to Jehan, there is at most one example of this use. Jehan's preference for use of the elided *ma* etc. therefore appears to be very strong, even if he may very occasionally have made use of *mon* etc. before feminine vowel-initial nouns.

In order to determine how unusual this feature is for the J-corpus, as well as for *Guillaume* and *Robert*, we must interrogate the assumption, made by both Morawski and Munk Olsen, that by the time the texts of the J-corpus appear to have been composed, around the turn of the fourteenth century, *mon* etc. and *ma* etc. were already being used interchangeably before vowel-initial feminine nouns, allowing Jehan to stand out as one who only uses one such type. A broader statistical approach will also allow us to support the suggestions of Marchello-Nizia *et al.* (2020) and claim that, even into the initial decades of the fourteenth-century, it was not unusual for a poet to make no use of *mon* etc. before a feminine vowel-initial noun.[106]

An examination of the P-corpus, containing all the texts of fr24432 with the exception of those contained within the J-corpus, reveals that of all 215 instances of *mon* etc. appearing before a vowel-initial noun, only 9 of these nouns are forms which, according to the DMF, are almost certainly feminine. These examples are given in Table 3.43.[107]

---

[106] Comparison with the F-, L- or R-corpora would be unhelpful here, given the dates of composition of these texts largely predate the texts of the J-corpus, even though on occasion versions of these texts appear in the same manuscripts as the texts of the J-corpus.

[107] It is notable that 5 of these 9 instances involve *ame*, which is one of the forms identified by Marchello-Nizia *et al.*(2020) as continuing to appear with elided *ma*, etc. towards the end of the fifteenth century. It is therefore possible that such forms may have been amongst the first vowel-initial feminine nouns to appear with *mon*, etc.,

Table 3.43 – examples of mon *etc. before unambiguously feminine vowel-initial nouns in the P-corpus.*

Text	Folio	Form	LH	KW	RH
Salhadin	32va	Verse		Son	espie[108] que moult amoit
Boivin	50rb	Verse	Que il voudra semer	son	orge
Bible	91ra	Verse	[…] ne plus fine,	Son	aime[109] pour biauté […]
Arbre	266rb	Verse	Mais por miex	mon	aaise avoir
Messe	321ra	Prose	[…] acorde son cors et	son	ame a Dieu […]
Lucidaires	372va	Prose	[…] qui fust bien de luy et	son	acointe […]
Lucidaires	373rb	Prose	[…] et de respoitier	son	ame ausint come […]
Lucidaires	375rb	Prose	[…] le sauvament de	mon	ame […]
Lucidaires	379va	Prose	[…] donne son cors au monde et	son	ame au dëable […]

At the same time, there are 151 examples of an elided form of *ma* etc. appearing before a feminine vowel-initial noun, meaning that across the P-corpus the preference appears to be for the avoidance of such an innovation. This is unlikely to be a result of scribal preference either, particularly in the manuscript's poetic texts, since to shift from *mon* etc. to *ma* etc. in this way would have an impact on the metre of verse lines.

Indeed, it would be tempting to hypothesise that scribal influence could have operated in the other direction since, as we argued in the first chapter, fr24432 was copied towards the middle of the fourteenth century, and certainly after 1345. It is therefore possible that scribes copying a text some decades after its initial composition could have chosen to change *ma* etc. to *mon* etc., particularly in prose texts, as linguistic practice had developed. However, none of the scribes in fr24432 appear to have attempted to make any consistent modifications to this feature within the texts that they copied, since Table 3.43 contains only 6 such examples from prose across the whole of fr24432, suggesting a remarkable consistency in both authorial and scribal practice in a preference for the elided forms of *ma* etc. before feminine vowel-initial nouns.

but only in part, with the process being completed more swiftly for other nouns. This is not the focus of our analysis here, however, and further work would be required to prove it.

[108] *espie*, 'spy.'

[109] As in *ame*.

That said, there is some small amount of evidence for the influence of scribal preference which comes from the fact that all the examples in *Lucidaires* appear within a few folios of each other, despite the text spanning some 58 folios, where numerous elided forms of *ma* etc. are also found. It is therefore unlikely that this results from intervention by the scribe who copied *Lucidaires* in fr24432, since the whole text is copied by this single scribe (hand 10). Rather, it would appear that this is perhaps evidence of variation in the scribe's exemplar, which itself may have been copied by a number of different scribes with different preferences, whereas the scribe of *Lucidaires* in fr24432 copied such forms as they found them. Scribal preference does therefore appear to have at least had some impact on the forms used in the manuscript, but the scribes in question are those who appeared further up the chain of transmission.

Given what appears to be a reasonably comparable usage of possessive adjectives across the diverse texts of the P-corpus, including both those which are similar in form and content to the texts of the J-corpus, and those which are significantly distinct, we are therefore drawn to the conclusion that Jehan's tendency not to use *mon* etc. before a feminine vowel-initial noun is not as surprising a feature as Morawski suggested. With limited exceptions, it appears likely that for the authors of all the texts found in fr24432, it was primarily an elided form of *ma* etc. that was preferred before a feminine vowel-initial noun. That is not to say that this feature is not worthy of observation in the J-corpus, but rather it reinforces Munk Olsen's warning that attribution cannot be made on this feature alone, since there were clearly other writers with similar preferences whose texts were circulating alongside the texts of the J-corpus.

Looking beyond the P-corpus of texts in fr24432, let us consider four verse texts included in the temporally broader B-corpus which appear roughly contemporary to when the texts of the J-corpus appear to have been composed, that is, the first quarter of the fourteenth century. These are *Le Dit du Prunier*, *Le Livre de la Passion*, *La Passion*, and *Le Roman de Fauvel*, all of which are written primarily in octosyllabic rhyming couplets and together create a

sub-corpus of some 62,000 tokens (the *B1-corpus*), roughly comparable in length to the J-corpus. The diatopic variation present in the B1-corpus helps us identify the extent to which the use of possessive adjectives observed in the J-corpus lines up with that observed across a broader range of dialect areas than that of the P-corpus, which itself is largely written in varieties akin to that found in the J-corpus, namely a western *scripta* with Picard and *francien* influences.[110]

Across the texts of the B1-corpus, the picture is even clearer than for the P-corpus. There are 47 instances of an elided form of *ma* etc. appearing before a feminine vowel-initial noun, and 55 instances of *mon* etc. appearing before a vowel-initial noun, all of which are either masculine or words for which the gender cannot immediately be determined. As such, these texts fit the trend observed across the J- and P-corpora in favour of the use of elided forms of *ma* etc. rather than *mon* etc. before feminine vowel-initial nouns, supporting the suggestion that Jehan's use of possessive adjectives was not unusual, even outside of his immediate dialect area.

Expanding the search to prose, if we take the four prose texts within the B-corpus which match the texts of the J-corpus in their estimated date of composition, we reach a sub-corpus of some 230,000 tokens (the *B2-corpus*), containing *La Chronique de Morée*, *Les Plaids de la sergenterie de Mortemer*, *Fouke le Fitz Waryn* (mixed poetry and prose), and *Les Mémoires ou La Vie de saint Louis*. These texts add a number of other *scriptae* to our comparison,[111] and in this case it is notable that, unlike for the B1-corpus of verse texts, in the B2-corpus there are 16 uses of *mon* etc. before a clearly feminine vowel-initial noun, with examples appearing in all four

[110] The presumed dialects of the authors of the first two texts of the B1-corpus, as determined by the team behind the BFM project (Guillot-Barbance *et al.* 2022) are both Picard, while the second two texts are Burgundian and Norman respectively, giving a good spread across what is now northern France.
[111] The dialects of the first of these texts has not been determined for the BFM, but the remaining texts have *scriptae* consistent with Norman, Anglo-Norman, and Champenois varieties respectively.

texts, out of 208 total uses of the adjective before a vowel-initial noun of either gender, whereas in 36 cases an elided form of *ma* etc. is used instead. While there are still more examples of *ma* etc. than *mon* etc. before a feminine vowel-initial noun, this rate of usage is much higher than seen in the J-, P-, or B1-corpora.

The preferences of later scribes appear not to have played a role in this choice, since all the transcriptions in the B2-corpus except for *La Chronique de Morée* are from manuscripts compiled within a decade or so of the texts' composition, and are therefore earlier than fr24432. Given that the B2-corpus differs from the prose texts of the P-corpus as well as the verse texts of the J- and B1-corpora, we are thus able to conclude only that, in the first decades of the fourteenth century, the use of *mon*, etc. before vowel-initial feminine nouns is variable in a way that is not strictly determined by text type or dialect area, but does at least appear to be rarer in poetry.

Returning to the question of the use of possessive adjectives in the J-corpus, we can therefore also conclude that, while Morawski and Munk Olsen were right to note that, at the time and place of Jehan's writing, it was indeed possible to place *mon* etc. before a feminine vowel-initial noun, it was still a lot less frequent than they assumed, and thus the absence of this feature in the J-corpus, while a unifying one, is not in any way surprising.

This analysis therefore does not assist directly in the collective attribution of the J-corpus, in that we have determined that variation in the usage of possessive adjectives in this period is not necessarily a good metric by which to determine authorship. That said, it does have methodological utility in determining the value that can be gained from the use of digital textual analysis to re-examine absolute claims made by pre-digital scholarship about the use of individual words. By analysing a broader array of texts than those considered by previous scholars, we have been able to come to a more nuanced and evidenced conclusion regarding

the linguistic preferences observed in the literature of the early fourteenth century. In turn, this has helped us clarify that for this period, non-use of *mon* etc. before vowel-initial feminine nouns is not an especially notable feature for a given author writing in early-fourteenth-century French, regardless of the dialect area from which he may hail.

3.9 Sources

Munk Olsen (1978: xxi–cxx) offers a comprehensive examination of the possible sources of the texts in the J-corpus, which I will summarise here, before turning to an examination of the contribution that can be made by digital textual analysis. For a number of Jehan's texts, it is possible to assert their sources confidently. The second tale within *Eaue et vergier*, for example, is either the direct or indirect descendant of the *Miserere* by the Recluse of Moiliens,[112] while *Chanoinnes*, *Juitel*, and *Merlin* are all based on tales found within the *Vie des peres*.[113] *Respon*, *Romme*, *Chevalier et escuier*, *Pecheresse*, and *Abesse* also share significant similarities with stories within the *Vie*, but are not close enough to be able to claim that they share a direct relationship. In addition to these texts, a further 11 tales appear to be original combinations of elements found elsewhere, existing in an established narrative tradition, with more or less significant modifications.[114] *Pommes*, *Florence*, and *Anelés* resemble pre-existing saints' lives, with modifications and additions that do not appear elsewhere, and to this we may also add that *Guillaume* and *Robert* also appear to be reworkings of existing textual traditions focussed on a single pious individual.[115] Kleist (1973: 53–55) also notes the presence of stories resembling *Respon*, the second tale of *Eaue et vergier*, and *Mescreant* in the *Ci nous dit*, a collection of very

[112] Munk Olsen 1978: lxxxv. For an edition, see van Hamel (1885).
[113] The first two of these connections were made by Gröber (1888–1902), and the third by Morawski (1939). No complete edition exists of the *Vie des peres*, but for the putative sources of *Chanoinnes*, see Chaurand (1971), *Juitel*, see Wolter (1879), and *Merlin*, see Méon (1823).
[114] *Deux chevaliers*, *Enfant rosti*, *Povre chevalier*, *Narbonne*, *Chevalier hermite*, *Enfant sauva mere*, the second tale in *Eaue et vergier*, *Pain*, *Mescreant*, *Beguine*, and *Sauveur*.
[115] *Pommes* shares similarities with *Ami et Amile*, an alexandrine version of which is edited in Neufang (1914). For further details on tales resembling *Florence*, see p. 253 and following, Wallensköld (1907), and Beaussart (2000). *Anelés* closely resembles *La Fille du comte de Pontieu*, edited by Brunel (1926). For more on the sources of *Guillaume* and *Robert*, see notes 56 and 66.

short summaries of didactic stories apparently intended as an *aide-mémoire* for sermon writers.[116] Only *Cordouanier* and *Buef* have no known analogues.

On the whole, then, Jehan appears to have operated within his existing textual context, and his originality comes not in the creation of new stories, but in his *remaniement* of existing ones, abridging longer tales and expanding on shorter ones, all with the intention of promoting his particular brand of narrative didacticism. The poet himself acknowledges having drawn inspiration from sources which he had encountered, but in ways which offer little direct evidence as to the texts' sources, and through claims which scholarship has refuted. Kleist (1973: 53) and Munk Olsen (1978: cxix) in particular note the presence of a number of claims regarding the written sources of the J-corpus, with Kleist arguing that 'Zwar weist Jehan häufig auf eine schriftliche Vorlage hin, aber solche Angaben sind formelhaft und besagen einfach, daß der Autor glaubt, sich in Übereinstimmung mit der schriftlichen Tradition zu befinden.' Similarly, Munk Olsen argues that any suggestion within Jehan's texts that his primary sources were written are little more than 'des clichés commodes,' suggesting that Jehan's sources for his poems were in fact primarily oral, but that this was a reality which the author attempted to conceal.

3.9.1 Claimed oral sources

Approaching this question of oral versus written sources through a digital lens can, however, be useful since, as we have seen, such an approach allows us to take a holistic view across the whole corpus. In this case, digital textual analysis allows us to comprehensively examine all of Jehan's references to his sources, and consider how credible they may be seen to be.

Our earlier examination has revealed that references to oral performance in Jehan's texts are particularly common, but it is notable that references to oral reception of his sources are

[116] On this, see our examination of *exempla* on p. 164.

much rarer. There are, in fact, only 2 instances in which Jehan suggests that material for his poems came to him by oral means. The first of these to appear is within *Anelés*, in which Jehan claims that '[…] moult l'ay oÿ conter / Que ja ne perira ce que Diex veult sauver' (vv. 555–556). This is of course nothing but an appeal to conventional oral wisdom, and therefore may offer some indication of Jehan's use of such popular didacticism when composing his poetry, but it does not give any indication as to how Jehan experienced the texts which are believed to have been the sources for his own poetry.

The only example of such a claim therefore comes in *Buef*, in which Jehan says of his tale that 'En sarmon l'oï dire, moult bien retenu l'ai' (v. 31). The specific nature of this claim, that Jehan encountered this story in a sermon, adds credibility to its possibility as a genuine indication of the poem's source material. Even if this is not the case, however, then at the very least it does foreground the supposedly religious origins of the tale, as well as signposting Jehan's awareness of the position of his poetry in an ongoing tradition of oral reception and performance. Ultimately, however, being the only reference anywhere in the corpus to suggest that the primary source for the text was oral, this example on its own is not enough to support the absoluteness of Kleist and Munk Olsen's claims regarding the oral nature of Jehan's sources.

Indeed, this sense that the texts of the J-corpus in fact do not foreground the apparent orality of their sources is reinforced by the prevalence of such claimed oral sources in certain of our reference corpora. The F-corpus, for example, is shorter than the J-corpus and yet contains some 6 instances of claimed oral sources, including the statement in *Boivin de Provins* (v. 2) in which the narrator, apparently the voice of the poet Jehan Bodel, claims that their story is one 'que j'oï conter a Douai.' The claimed oral sources of the *Lais* of the L-corpus, meanwhile,

are well studied, with the poet herself making claiming in the *Prologue* that she has chosen to engage with oral sources in explicit contrast to those who use written sources.[117]

3.9.2 Claimed written sources

Reading

Turning our attention away from the avenues suggested by previous scholarship, we find that while references to texts experienced through oral performance are surprisingly limited in the J-corpus, that he makes a good deal more claims to have gathered at least some of his material from written media. For example, across the poems of the J-corpus, we find 3 uses of forms of the phrase *trover lisant* to suggest the written nature of his sources, which are given in Table 3.44.[118]

Table 3.44 – uses of the phrase trouver lisant *in the J-corpus.*

Text	Line	LH	KW
Enfant rosti	45	Huit jors devant Noël, si com	trouvons lisant
Narbonne	18	A l'escole fu mis, si com le	truis lisant
Enfant sauva mere	29	Il fu .i. riches hons, si com	trouvons lisant

All three examples appear somewhat out of the blue, in sections of the text that are setting up the narrative but make no other efforts to identify their sources. As a result, while they already provide a greater number of references to written sources than to oral ones, they are not sufficient on their own to confidently refute Munk Olsen's hypothesis that all such references are simply 'des clichés commodes,' included to add a sense of authority to Jehan's words, rather than a specific reference to his sources. Indeed, we may note the lack of any more specific detail surrounding these phrases, as well as the fact that all three examples appear in apparently fixed constructions lasting a single hemistich, with *lisant* at the end of the

[117] On this, see for example Bruckner (2005).

[118] This is seemingly not a phrase unique to Jehan, with 15 examples occurring in the B- corpus, although no such examples appear anywhere in the P-, F-, or L-corpora. Indeed, no uses of the verb *lire* appear anywhere in the F- or L-corpora, other than to refer to elements of writing within the narrative.

line, suggesting possible use simply as means of filling a line with an 'easy' rhyme. The use of the first-person plural in two of these examples further reinforces this suggestion, since it adds a sense of generality, indicating that one could find such references in written form, but not necessarily suggesting that Jehan himself has done so.

The verb *lire* also appears in the first-person plural in *Guillaume* (f. 2^{vb}), this time without a form of *trouver*, in 'si comme nous lissons.' This phrase appears in a similar context to those observed in the J-corpus, and helps us to reinforce our attribution of this text to Jehan, while also reinforcing the suggestion that such phrases, whether true or not, are used predominantly as a way of adding authority to a given text while also filling an entire hemistich and in this case providing a rhyme.

Finally, we may also argue that the verb *trouver* alone is a term more generally associated with reading, through association with its use in *trouver lisant*. A number of examples of this use within the J-corpus will appear alongside those given for other reasons in the following sections, but it is worth for now noting a passing reference in *Chevalier hermite* that 'Nous trouvons que jadis estoit .i. chevaliers' (v. 17), arguably an oblique reference to writing, which suggests again the use of such references as a means of asserting textual authority.

Writing

By looking at references in the J-corpus, as well as in *Guillaume* and *Robert*, to Jehan's sources having been written down, we find further examples which create a body of evidence against the absolute nature of Kleist and Munk Olsen's claims regarding Jehan's sources. As an initial point of comparison, we note that there are only two instances within the F-corpus in which the narrator refers to a written source for their text, which in one instance is placed on an equal footing with oral sources. In the L-corpus, meanwhile, as we have seen, Marie de France ostensibly shuns written sources in favour of oral ones. There are however two

occasions when Marie makes reference to written sources for her text, namely the claim in *Guigemar* that she will tell her tale 'sulunc la letre e l'escriture' (v. 23), and the suggestion in *Chievrefueil* that Marie's sources span both oral and written media, since 'Plusur le m'unt cunté e dit / e jeo l'ai trové en escrit' (vv. 5–6).

By contrast, across the J-corpus, *Guillaume*, and *Robert*, there are some 17 examples concerning writing, included in Table 3.45, which give some suggestion that all or part of a given text is drawn from a written source, either directly or through more oblique means.[119] These examples appear across 9 texts of the J-corpus, as well as both *Guillaume* and *Robert*, therefore suggesting a degree of consistency across the J-corpus, while again reinforcing the strength of our attribution of the latter two texts to Jehan.

Table 3.45 – references to written sources in the texts of the J-corpus, as well as Guillaume *and* Robert.

Text	Line/Folio	LH	KW	RH
Chanoinnes	1	L'	escripture	tesmoingne et nous fait asavoir
Romme	81	Elle fist tant de biens si com dit li	escris	
Romme	103	Et de ces .ii. pechiez si com dit li	escris	
Deux chevaliers	13	Il nous est tesmoignié en la sainte	Escripture	
Eaue et vergier	9	La divine	Escripture	nous en fait mencion
Pain	13	Ainsinc com nous trouvons en la sainte	Escripture	
Mescreant	21	En la *Vie des peres*	escript	en parchemin
Florence	20	Que l'	escripture	dit de ce ne doutez mie
Buef	87	Grant penitence en firent si com dit li	escris	
Sauveur[120]	4	Ainsi est il	escript	en la divinité.
Sauveur	11	Or advent une fois, si come dit l'	escrips	
Guillaume	1ra	L'	escripture	tesmoingne que qui vault hault monter
Guillaume	1ra	Et c'est	escript	c'un riche qui ne veult dire rien
Guillaume	1rb	Le philosophie dist – c'est	escript	en maint lieus –
Guillaume	2vb	Que l'	escripture	dit c'om doit pour Dieu laissier
Guillaume	8va	Que je croi qu'il n'est clers qui le peüst	escrire	
Robert	203rb	Ainssi comme l'en trueve en pluseurs lieus	escrist	

[119] This is in addition to 5 references to writing within the narrative and 2 references to Holy Scripture without discussion of how its content has influenced the narrative.

[120] The line immediately preceding this one reads: 'Les sages clers le tiennent en grant auctorité.' While this is not an explicit reference to writing, it certainly gives a further indication that Jehan believed, or at least wished to claim, that his tales stemmed from a place of textual authority.

Further evidence for the collective attribution of *Guillaume* and *Robert* to Jehan through this feature comes from the similarities in the structures they contain. *Guillaume* shares 'L'escripture tesmoigne' with *Chanoinnes* and 'Que l'escripture dit' with *Florence*, while 'Ainssi comme l'en trueve' in *Robert* is reminiscent of 'Ainsinc com nous trouvons' in *Pain*, and indeed the two texts share a similar structure, with 'c'est escript en maint lieus' in *Guillaume* resembling 'en pluseus lieus escrist' in *Robert*. The narrator's claim in *Guillaume*, that the eponymous king's suffering was so great 'Que je croi qu'il n'est clers qui le peüst escrire,' also lines up with one found in direct discourse in *Deux chevaliers*, that the torments of hell are such that 'Cuer nel porroit penser ne main de clerc escrire' (v. 124).[121]

Notably, none of the examples in Table 3.45 appear within texts that also contain examples of *lire*, thus expanding the total number of texts in the J-corpus containing references to reading written sources to 13, over half of the corpus, which reinforces the apparent importance of such references for Jehan.

In addition to this, two features of these examples stand out. Firstly, as has already been observed for *trouver lisant*, 8 of the 17 examples appear at the end of the line. Most of these line-final examples occur within one of two construction types, both of which span a single hemistich, which again suggests at least some use of the reference as a means of filling out the poetic metre. The most common of these, *si com dit li escris* (4 examples), could itself be reasonably easily dismissed as hemistich-length phrase which can simply be tagged on to a line to provide an appropriate rhyme, while also adding a sense of textual authority. However, for the second structure which appears at the end of a line, *en la sainte Escripture* (2

[121] I make no claim as to whether this phrase in *Guillaume* is an oblique self-reference by Jehan, since I do not wish to engage in the semantic debate around whether Jehan was a *clerc* (as Munk Olsen (1978: xi–xii) claims) or a *jongleur* (as Faral (1910: 176) posits).

examples), the question of function is a little less straightforwardly about filling a line, since both appear as part of wider phrases which emphasise the reference to a narrative source.

The possible function of these latter structures is made clearer through the examination of a second feature to stand out within these examples, namely the claim that the written sources from which the texts allegedly emanate are in some way *sainte* (2 examples, both in the repeating structures discussed above) or *divine* (2). Interestingly, the two references to Holy Scripture elsewhere in the J-corpus also make use of the descriptor *divine*,[122] with one such example appearing within *Pain*, which itself contains a description of Jehan's supposed source for the poem as *la sainte Escripture*. In the examples involving *divine*, the references are to general moralising claims rather than to a narrative source, with no mention made of the exact source of the claim. From the similarities of these structures to those involving *sainte*, we therefore get the impression that, through his use of such phrases, Jehan was similarly intent on claiming that *Deux chevaliers* and *Pain* were drawn directly from Biblical or apocryphal sources, without offering any evidence that this is the case. He uses the terms *sainte* and *divine* to add authority, rather than clarity, to his supposed narrative and didactic sources, and his references to writing in this way also act as little more than markers of this authority, rather than invitations for an audience to go and seek out the sources themselves.

This aligns with the four examples found towards the beginning of *Guillaume*, in which the narrator refers to written sources in order to reinforce a number of pieces of conventional wisdom which he presents. In these cases the written sources are not said to be *divine* or *sainte*, but the religious and didactic content of the argument presented, as well as the reference to *philosophie*, reinforces the suggestion that Jehan's inclusion of ideas of writing was often as a means of adding authority to the didactic component of his poems.

[122] *Enfant sauva Mere* (v. 14) and *Pain* (v. 7).

This is not necessarily true for all of Jehan's poems, however. In contrast to the general nature of these examples, and in greater contrast to Kleist and Munk Olsen's hypotheses that Jehan's references to writing were simply rhetorical flourishes, a surprisingly specific example is found in *Mescreant*. The text tells of a disbeliever who attempts to prove the futility of the Eucharist by feeding the Host to his dog, which refuses to take it and devours its owner instead.[123] Jehan claims that the source for his version of this text is the *Vie des peres*,[124] itself apparently 'escript en parchemin.' Again, this is clearly an appeal to authority, but it would seem premature in this case to so easily dismiss this example, as Munk Olsen (1978: cxix) does, as a reference which does not 'correspondre à [une] réalité.'

Indeed, the reference to material transmission is sufficiently specific as to warrant a closer examination of the *Mescreant*'s possible source. No clear ancestor of *Mescreant* appears within the surviving versions of the *Vie des peres*, with Munk Olsen himself finding only a tenuous connection to a tale within Optatus' fourth-century *De schismate donatistarum*, in which heretics are punished for having fed the Host to their dogs.[125] Both from the differences between the content of the tales, and Jehan's focus on vernacular literature, it is clear that he did not work directly from the *De schismate donatistarum*, instead he most likely borrowed the tale from a source much later in the chain of the tale's transmission, if the two are indeed connected.

---

[123] This tale resembles, but is far from identical to, Tubach's (1969) tale type 2641, *Host adored by dog*, in which a Jew offers the Host to a dog, but the dog bows before it. Similarities are also found to 1718, *Dogs tear apart excommunicated woman*, a fate also sometimes given to a false priest, as well as 3776, *Pig, scrupulous*, in which a pig, or sometimes a dog, refuses to take bread from the hand of an excommunicate. The *mescreant* at the heart of the tale, we should note, appears to be a Christian who does not believe in the power of the Host, rather than a Jew. It would not be a significant jump, however, to suggest that Jehan's source(s) may have focused on Jewish Host desecration. For more on this (but nothing resembling our present tale), see Rubin (1999).

[124] Two possible further references to the *Vie des peres* are made in *Eaue et vergier*, which references 'les sains peres' (v. 98) as the source of the poem's second tale, and *Sauveur*, which claims its source as the '*Vie des sains*' (v. 5). In both cases is unclear whether or not these references are to the same sequence of texts as the *Vie des peres* referenced in *Mescreant* (whether that be the texts known today as the *Vie des peres* or some other sequence).

[125] Book ii, chapter xix: *Eucharistia a Donatistis canibus data, ampulla chrismatis in praeceptis projecta aliaque eorum sacrilegia*. For an edition, see Migne (1844–1865, xi, cols. 972–974).

Since no text survives which could offer such a link, it is therefore not unreasonable to suggest that Jehan's claim is not false, and that he did indeed draw his inspiration from a source, now lost, which told a similar story to *Mescreant* in French, and which may itself have been included within versions of the *Vie des peres*, or another sequence known by the same name. After all, the fact that the surviving texts now known as the *Vie des peres* do not include the text which Jehan claims as his source does not mean that his claim is not truthful.

Similarly, while we cannot determine with any certainty whether Jehan's interaction with this putative text would have been through the written or aural medium, and even if we take a less absolute view of Munk Olsen's perspective, assuming that Jehan's work was primarily completed from oral sources, the reference to parchment in *Mescreant* means that it is certainly possible that he saw this text in the flesh, or indeed other texts which he took as his sources. Turning to the B-corpus as a source of comparison, we find that the phrase 'en parchemin' and its orthographic variants appears on 15 occasions, with only 3 referring to the sources of a text. Particularly comparable to Jehan's use is the structure found in the prologue to the *Lai de Tyolet*,[126] in which the narrator claims that stories of King Arthur's court 'mises estoient en latin / e en escrit em parchemin, / por ce qu'encor tel tens seroit / que l'en volentiers les orroit' (vv. 29–32). In this case, the narrator claims an intervening role played by the Bretons, who '[...] en firent lais plusors, / si con dïent nos ancessors' (vv. 35–36). It is the Breton *lai* that the narrator claims to relate, and thus his claimed sources are oral, despite the reference to written material. Jehan's claim may therefore be similar, presenting a belief that written material appeared further up the chain of the story's chain of transmission, even if the source from which Jehan drew his inspiration was aural. Given the lack of any reference to intervening sources in *Mescreant*, unlike in *Tyolet*, our conclusion for the text's source must be more tentative, and allow for the possibility either that Jehan's primary source was written, or

[126] For an edition of this text, see Walter *et al.* (2018).

that he had at least encountered a written version of the story. We must, of course, not assume that Jehan would only have encountered a given tale once, and in one particular medium: it is certainly plausible that he could in fact have come across the same tales on a number of occasions and in different media.

Brémond, Le Goff, and Schmitt (1982: 63–64) offer some useful context here, namely that as the fourteenth century progressed, a great deal of the reception of *exemplum*-type narratives, which Jehan may have drawn inspiration from, was through the written, rather than the oral medium, in the form of moralised tales intended to be read, such as Nicolas Bozon's *Contes moralisés* (ca. 1320),[127] which take a tale akin to an *exemplum* as their starting point, but offer additional moral explanation surrounding it, as Jehan does.[128] While Jehan, if indeed he did write at the beginning of the fourteenth century, would be early in this process of change in the *exemplum* from an oral to a written tradition, this possibility suggests that for the case of *Sauveur*, it is at least feasible that Jehan encountered the tale which inspired it through a written text which has now been lost.

We therefore ought at least to entertain the possible veracity of Jehan's claims as to the written nature of certain of his sources, although we ought at the same time to be more cautious with other examples. The prime instance of this is to be found in *Florence*, in which Jehan claims on two separate occasions to have sourced his tale from the chronicles of Rome. In the first of these, he compares them in content to the chronicles of Saint-Denis:[129]

> Douce gent, es croniques de Saint-Denis en France,
> Voit on moult de merveilles ; mais sachez sanz doutance
> Celles de Romme sont de trop plus grant sustance. (*Florence*, vv. 5–7)

[127] For an edition, see Toulmin Smith & Meyer (1889).
[128] Indeed, we may choose to see *conte moralisé* as the most appropriate generic descriptor for Jehan's texts.
[129] Cholette (2022: 102) erroneously suggests that it is the Saint-Denis chronicles that Jehan is positing as his source, a claim which does not align with the content of this stanza, nor with the Roman flavour of the text as a whole.

Such a reference, as well as the comparison given, is particularly effective as a means of reinforcing textual authority, since Jehan's claims to know the chronicles of both Saint-Denis and Rome paints an image of the author-narrator as particularly well-read. Beaussart (2000: 6) interprets this as being a reference to the version of the tale in the *Gesta Romanorum*,[130] but as we have seen, Jehan's lack of engagement with Latin material suggests that this is unlikely to have been his direct source. Further, as Beaussart himself acknowledges, the fact that the text does not contain a narrative element in which Florence is accused of adultery with her brother-in-law and briefly condemned to death by her husband suggests at the very least that Jehan 'ne s'inspire pas exclusivement des *Gesta*' (2000: 6), such that the text may be seen to be closer to the thirteenth-century *Chanson de Florence de Romme* (2000: 8).[131]

A second reference to the chronicles is revealing of a broader literary awareness, in which Jehan shores up the claim that his version is superior to others. At the moment when Florence's abusers arrive at Beaurepaire to meet their judgement, Jehan asserts:

> Douce gent es croniques de Romme sont trouvees
> Les paroles qui sont ci de par moi contees,
> Mais .i. rommans en est ou en a ajoustees
> Granz bourdes qui n'i doivent pas estre recordees (*Florence*, vv. 693–696)

On this occasion, Jehan offers a criticism of an alternative version of the same tale, itself possibly also written down, which supposedly includes information which Jehan considers inappropriate, or simply extraneous. If Jehan is making a specific criticism of the content of a version of the text that he read, it is unclear exactly which this might be, since none of the surviving versions of the tale differ significantly from *Florence* at this particular moment. Given Jehan's references to the *Vie des peres* elsewhere in his corpus, however, it is at least also possible that he chose to critique a version of the tale linked to the one which survives in this

[130] For a recent edition, see Weiske (1991–1992).
[131] For an edition of this text, see Wallensköld (1907–1909). Beaussart also examines *Florence* alongside the thirteenth-century *De l'empeeris qui garda sa chastee* by Gautier de Coincy, edited in Méon (1823: 1–128).

collection,[132] which also contains the motif of accused adultery while also omitting further details present in Jehan's version, including for example, the murder of a child in Florence's care (Beaussart 2000: 5, n. 5).

Whether or not we believe Jehan's claim to have encountered this 'rommans' in particular, we may interpret this reference as an attempt by Jehan to see off his literary competitors who may have been telling similar stories. It is also not inconceivable that, regardless of the texts to which he refers, Jehan consulted at least two versions of the tale in his preparation for writing *Florence* — one which he used as his main source and one which he rejected — which would reinforce our rejection of the absolute claim that Jehan had a single oral source for each of his texts.

If we consider Jehan's references to the chronicles of Rome in light of this apparent awareness of multiple versions of the same story, our reading of such claims can be nuanced somewhat. Rather than reading it as a claim that his texts were drawn directly from the chronicles, or perhaps the *Gesta Romanorum*, one might suggest that Jehan here is citing his sources second-hand, and that his inherited belief, genuinely held or otherwise, was that the chronicles were the ultimate source of the story. Jehan is not claiming to have found the stories himself, as evidenced by the passive 'sont trouvees,' but he sees himself as participating in a textual tradition that stretches back to the chronicles, with his poems offering a synthesis and *remaniement* of various later versions of the tale which he had encountered, perhaps in both written and oral media.

3.9.3 A possible contradiction

This more nuanced perspective, that Jehan most likely drew inspiration from a number of different sources for his tales, which may have been both oral and written, allows us also to

[132] On which see Wallensköld (1907).

take a more considered approach to an apparent contradiction observed by Munk Olsen (1978: cxix) in the sources claimed within *Buef*. As we have seen, at the beginning of the text, we find the claim that 'En sarmon l'oï dire, moult bien retenu l'ai' (v. 31), which Munk Olsen argues is 'certainement une description exacte de sa méthode de travail,' and which Cholette (2022: 102) suggests is evidence of the poem presenting its sources as exclusively oral. A little later on, however, we find the structure 'si con dit li escris' (v. 87), as given above in Table 3.45.

Given our previous conclusions about this latter phrase, it would be easy to dismiss it, as Munk Olsen does, as one of the 'clichés commodes qu'il faut se garder de prendre en sérieux' (1978: cxix), allowing us to take Jehan's claim regarding the oral nature of his source at face value. That said, the overwhelming abundance of references to written sources found throughout the rest of the J-corpus, as well as for *Guillaume* and *Robert*, cast doubt on the strength of Munk Olsen's claim that there is no way Jehan could ever have encountered his sources through a textual medium. Instead, if we assume that Jehan's poem had multiple sources, then it is entirely plausible that he could have both heard the tale told in a sermon, and come across it in writing. In fact, this appears all the more plausible an interpretation when we consider that, as for *Florence*, *Guillaume*, and *Robert*, Jehan's tendency is far more often to shorten texts, rather than lengthen them, and thus a short sermonic *exemplum* is unlikely to have been the sole source for *Buef*, itself the second longest text in the J-corpus.

Overall, then, a more thorough digital textual analysis than has previously been possible has allowed us to argue that the claim that Jehan only encountered his texts through oral sources is at the very least excessively inflexible. Ultimately, there is no good reason to ignore the examples of specific references to written sources which appear within his texts, even if we are to be wary of taking their claims at face value. From the perspective of the attribution of authorship, we therefore observe a unity across the J-corpus, as well as for *Guillaume* and

Robert, but not in the way that previous scholarship seems to have imagined. Jehan's texts are united in the frequent use of references to written sources, and the widespread absence of references to oral ones, with the poet making a conscious effort to appeal to the authority of written text as a means of strengthening his own arguments. As we have seen, this differs from the approaches presented in the F- and L-corpora, where references to claimed oral sources are more common, and thus while references to the written sources of a text are far from unique to Jehan, we are nevertheless able to add this evidence to the overwhelming sense of unity which we have observed in Jehan's texts up to this point, facilitating their collective attribution.

3.10 Conclusion

As we have seen, there is no silver bullet within either the digital or traditional methods we have discussed that proves Jehan wrote these and only these poems. Our best chance of accurately attributing these texts comes instead from a holistic examination of as many types of evidence as possible. Munk Olsen (1963b: 163), who is so critical of Morawski's willingness to assign or not assign texts to Jehan based on a single feature, nevertheless acknowledges the value of these individual strands of evidence when taken as a whole:

> [U]n critère isolé ne prouve rien ou plutôt peut prouver n'importe quoi, ce qui est encore pire. La seule méthode adéquate est d'établir une convergence des ressemblances entre les […] œuvres.

This is exactly what I have attempted to achieve in this chapter, in which digital and statistical methods of analysis unavailable to previous generations of scholarship have allowed us to widen the variety of evidence for this divergence in Jehan's texts. The presence or absence of a particular feature in any text cannot, in any scientific manner, prove that it shares authorship with another. This is even more the case for medieval texts, given that they exist within a climate of textual borrowing, influence, and reuse. Yet by taking a global,

holistic view of different strands of analysis, we are able to gain a much clearer picture than any one strand might suggest.

Through the examination carried out in this chapter, we have found that in the case of Jehan's texts, as observed for example in his versificatory practices, significant variability between the texts of the corpus can, without the right analysis, mask the extent to which the texts are still collectively different from others. We have also discovered that a given author will not differ from those in his literary milieu in every metric which we may consider, as with our examination of Jehan's use of language. This is not to say that such metrics should never be considered, rather it reinforces the value of a holistic approach to authorial attribution which allows for the possibility of overlap with other corpora, providing that the overall picture points consistently in a particular direction. Doing so has allowed us to assert with greater confidence that the texts of the J-corpus can all be comfortably attributed to Jehan, and it has also facilitated the attribution of a further two texts to the same author, namely *Guillaume* and *Robert*, not based on any one point of similarity, but rather as a result of the holistic examination of a large amount of evidence. As Cholette has suggested, the texts do indeed share a 'poétique commune' (2022, 95), but this is not, as he argues, solely due to the fact that the texts are members of the same generic group.

While it can be useful to draw conclusions from large-scale data on the basis purely of statistical trends, our examination of the generic descriptors appearing within and alongside Jehan's texts, as well as the terms of address he uses for his audience, has also revealed the value of the use of such techniques in the context of a broader understanding of the function and content of the source material as a whole. After all, the significance of a particular feature cannot be determined only in isolation, and can only be reached after a careful reading of a text and close engagement with its context.

Beyond the question of Jehan's authorship, we have seen that the methods of analysis used in this chapter are broadly applicable to a wide array of Old French texts. Through our examination of the F-, L-, and R-corpora, we have observed, independently of the J-corpus, that these corpora of more frequently studied texts display notable areas of coherence alongside a good deal of variation. The methods presented in this chapter could therefore also lend themselves to a comparative evaluation of the unattributed *fabliaux*, for example, as well as the texts in Rutebeuf's *œuvre* for which attribution is more tentative.[133] Even in cases when attribution is more assured, as in the *Lais* of Marie de France, a fuller analysis using the methods used here could allow for fruitful comparison between these texts and other *lais* which are not attributed to Marie. Outside of the texts of these reference corpora, it is also possible to foresee these methods having utility when applied to some of the most difficult questions of authorship surrounding texts in Old French, such as that of the *Lai d'Aristote*,[134] as well as several of the more peripheral texts attributed to Chrétien de Troyes.[135]

Analysis of these reference corpora, when combined with the broader P- and B-corpora, has allowed us to come to new conclusions of value to historical linguistic studies surrounding the extent to which a given feature, such as the analogical use of *mon* etc. before a feminine vowel-initial noun, may be considered typical for a given temporal and dialectal context. Most importantly for questions of authorship, it has also allowed us to tackle the thorny question of whether a shared feature of a corpus is a result of shared authorship or a feature more commonly found across a given genre. The F- and P1a-corpora have revealed that such common features do exist across the works of different authors which fall within a single

---

[133] In his 1989–1990 edition of Rutebeuf's texts, for example, Michel Zink places doubt on the attribution of both *La Vie du monde* and *Les neuf joies de Notre Dame*. An analysis of the kind that we have carried out here for Jehan's texts could therefore shed further light on the matter.

[134] See Corbellari (2004) and Zufferey (2004) for a combined study which follows many of the same lines of inquiry as followed by Morawski (1939, 1940) and Munk Olsen (1963a, 1963b, 1978) for Jehan's texts, but which does not venture into the methods presented in this chapter.

[135] See, for example, the possible attribution to Chrétien of the longer version of the tale found in *Guillaume*, discussed in note 66.

generic category, and yet the R-corpus has also shown that such common features are also possible when the works of a single author span multiple genres. For this reason, it is necessary for future studies of authorship using these methods to always consider a similarly broad selection of reference corpora, to determine when a shared feature may genuinely be an indicator of shared authorship.

In the case of *dits* written in alexandrine monorhyme quatrains, we have seen that form and genre play a seemingly significant role in the use of features such as feminine rhymes, as well as choices of discourse onset, and yet significant variability within a single genre is possible, and individual authors are able to make their preferences felt within their texts. Similarly, features such as the rates of use of words at the beginnings of lines, as well as associated rhymes and admixtures, appear to be more consistently the result of authorial preference.

To close this chapter let us end, as we began, with an examination of the question of why, if at all, one would consider attempting to attribute a medieval text to a specific author. The evidence for attributing medieval authorship will almost always be tentative, yet whether or not there was in fact a man called Jehan de Saint-Quentin who wrote these texts, it tells us a lot about the texts themselves that scholarship, and potentially also some earlier readers, have believed that they were all the work of the same individual, despite limited self-attribution to support that claim. Throughout the chapter, we have considered how the texts which we have attributed to Jehan display both unity as a group, and distinctiveness from other texts. We have done this as a means of determining shared authorship, but these features may also be fruitfully considered from a more literary perspective. Doing so allows us to encounter Jehan's *œuvre* as a network of interconnected texts, each of which contains features that call to mind those of another. Jehan's didacticism, his attempts to ease the sins of mankind, and his desire simply to tell an engaging story all come to the fore even more when we read his texts as a united collection, rather than when a given text is encountered in isolation.

As Jehan so often does in his own texts, we end this chapter with a stanza of verse in the form of a prayer, this time generated by ChatGPT (OpenAI 2022a), after having been given a selection of Jehan's epilogues as inspiration:

> Or prions humblement le Seigneur souverain
> Qu'il eclaire nos pas sur son divin chemin.
> En sa grace infinie, qu'il guide notre main
> Vers un destin certain en son royaume serein.

Emerging technologies such as ChatGPT reveal more than ever that the unique identity of an author is a fragile notion which can be so easily be broken by imitation and influence, and yet an author is not a Large Language Model. Digital and statistical methods of analysis have a significant role to play in the identification and attribution of authorship, but such an approach can only be effective when used in a humanistic way, combining scientific techniques with literary understanding.

Our analysis here has been far from exhaustive, but it has at least added to the voices supporting the attribution of the J-corpus to Jehan, as well as suggesting the novel inclusion of *Guillaume* and *Robert* into his *œuvre*. As we have done so, the chapter has displayed how much digital and statistical methods and approaches have to offer to questions of authorial attribution, when used in combination with more traditional approaches.

Conclusion

Across the three chapters of this thesis, digital methods of analysis have been used in partnership with more conventional methods. In Chapter 1, through a full material, structural, and palaeographical analysis of fr24432, this partnership was equal, with quantitative approaches brought to bear on the codex alongside more traditional qualitative methods, depending on the requirements of the details under examination. We distinguished the manuscript's 11 main scribal hands through first qualitative and then quantitative approaches, before turning to detailed examination of the manuscript's quire structure, the surviving evidence of pricking and ruling, the near-contemporary foliation, the *litterae notabiliores*, and the table of contents. From the evidence of complexity present in the manuscript, we concluded that fr24432 began life as three distinct manuscript units, which were created and then subsequently recompiled, all before the codex left its original place of production. It was through the combination of these qualitative and quantitative data that we were able to determine the order in which much of this processes of change took place, including the moments at which the original base manuscript had material added and removed, with the most significant modification being the later inclusion of an additional grouping of quires, containing the prose text *Lucidaires*.

Drawing from both traditional and digital perspectives, in Chapter 1 I have therefore been able to reach new conclusions regarding the production of fr24432. This mix of approaches was equally valuable in our attempts to localise and date the manuscript, such that despite its material and linguistic complexity, it is possible place fr24432 just to the north of Paris, at some time close to the middle of the fourteenth century. We have also been able to trace what little can be known of its provenance, both through external sources and through the clues left by those who have interacted with the codex across the centuries.

Following our opening analysis in which digital and conventional analyses were placed on an even footing, Chapter 2 took a digital-first approach, examining the use of scribal abbreviation within fr24432. We began the chapter from a more theoretical perspective, considering the place of abbreviation within medieval writing practices, and the previous attempts that have been made by scholarship to account for its use. Through this, I argued for the value of an approach to manuscript abbreviation that is more concerned with function than form, since this allows us to better account for the realities of how medieval scribes used abbreviation. Having established the theoretical basis for our analysis, we then turned to methodological considerations, stepping away from the methods and approaches used by previous generations of scholarship, in favour of a method which makes use of digital and statistical techniques to identify trends and shifts in the use of abbreviation within a given manuscript. Applying this method to fr24432, it was first possible to reinforce conclusions about the scribal structure of fr24432 that were made in Chapter 1, using the abbreviatory data to confirm many of the moments at which one scribe hands over to another. We were also able to conclude that scribal variation is by far the most significant factor determining abbreviation usage, and that certain non-scribal features play a less significant role, such as the position of a candidate for abbreviation relative to codicological features such as the beginning or end of a line, column, page, or quire, which for the most part have a limited impact. That said, there are other non-scribal features which are associated with more significant differences in the use of abbreviation, including whether a text is written in verse or prose, as well as the exemplar manuscript from which it is copied. Further comparative work is required to identify whether these trends are applicable across medieval French manuscript text more broadly, but the analysis carried out in Chapter 2 presents an approach that can be usefully applied to a number of different contexts, and provides a starting point for a more expansive study on the use of abbreviation in French manuscripts.

Finally, Chapter 3 presented a case study displaying how methods of digital and statistical analysis can be applied to questions of uncertain authorship even after a body of more conventional scholarship has built up. In particular, examining the case of Jehan de Saint-Quentin, data-driven approaches allowed us to affirm the attribution of those texts previously attributed to Jehan through concrete, reproducible results, while also adding two further texts to his *œuvre*, namely *Le dit de Guillaume d'Engleterre* and *Le dit de Robert le Deable*.

As part of the analysis present in this chapter, we examined Jehan's texts both individually and collectively, as well as through comparison against a number of reference corpora and sub-corpora, ranging from those which appear alongside many of Jehan's poems in fr24432 (the P-corpus), to those which exist as a distinct generic or authorial group (the F- and L-, and R-corpora), as well as a much broader corpus combining texts with a variety of spatial and temporal origins (the B-corpus). The features considered ranged from more formal ones such as versification, and shared vocabulary and structures, to textual features including details of claimed performance, genre, and sources, as well as to linguistic features. In this case, it was the holistic consideration of the whole analysis that ultimately allowed us to conclude that the texts of the reference corpus and our two hypothesised additions are all the work of Jehan, thus enabling us to contribute to a broader understanding of this otherwise obscure poet.

The primary methodological conclusion of the thesis as a whole is that the application of digital and statistical methods has much to offer the field of manuscript studies, but that these methods are at their most effective when they are developed in a bespoke way for the codex or codices under consideration. Every manuscript is a unique object, and deserves to be treated as such, and our understanding of manuscript culture will be vastly richer for an approach that does not seek to rigidly apply the same methods of analysis across vastly diverse objects. The methods which we have applied here, which have been chosen on the basis of an in-depth engagement with the manuscript and its texts, are not identical to those which would

be most useful elsewhere. As handwritten text recognition technologies become more widespread, and the quantity of digitised manuscript text increases, we must also remember that the encoded codex, much like the physical one, must be engaged with on its own terms.

Bernhard Bischoff notably feared that the advent of statistical and numerical techniques meant that manuscript studies, and palaeography in particular, would become 'an art of measurement,' and cease to be an 'art of seeing and comprehending.'[1] As we have seen, there is a great deal of value to be found in measurement, but this does not mean that we must abandon the humanistic understanding that digital approaches alone cannot capture. Manuscript studies and textual analysis must remain arts of comprehending, but this can be achieved through measurement, just as well as through seeing and reading.

By way of an *envoi*, let us consider the following stanza, found within fr24432's copy of *La vision de saint Paul*:

> Saint Pol a regardé en cele grant dolor ;
> Ne pot goute vëoir por cele tenebror.
> "Saint Michiel," dist saint Pol, "por Deu le Creator,
> Enlumine mes iex de ta grant resplendor,
> Que je puisse vëoir dedens cele oscuror." (f. 95v)

At first glance, a manuscript with as much material and textual complexity as fr24432 can appear as obscure as the abyss which Saint Paul peers down into. Each method of analysis which we have applied to it across this thesis, be it digital, statistical, codicological or otherwise, has added a little more light to the overall view. It would be far from the truth to say that we have seen all there is to see in this manuscript, and some of its darker corners may forever remain concealed, but this combination of diverse approaches has, like Saint Michael, illuminated our eyes, and made Bischoff's 'art of seeing' that little bit more assured.

[1] Cited and translated in Derolez (2003: 8).

Bibliography

Manuscripts

Berlin, Staatsbibliothek und Preussischer Kulturbesitz, Hamilton, 257
Bern, Burgerbibliothek, 354
Besançon, Bibliothèque municipale, 588
Grenoble, Bibliothèque municipale, 871
Hildesheim, Dombibliothek, St Godehard MS 1
London, British Library, Additional, 15606
Oxford, Merton College Library, MS 249
Paris, Bibliothèque nationale de France, Arsenal, 2115
Paris, Bibliothèque nationale de France, Arsenal, 3525
Paris, Bibliothèque nationale de France, Arsenal, 4629
Paris, Bibliothèque nationale de France, fonds français, 834
Paris, Bibliothèque nationale de France, fonds français, 837
Paris, Bibliothèque nationale de France, fonds français, 1593
Paris, Bibliothèque nationale de France, fonds français, 1630
Paris, Bibliothèque nationale de France, fonds français, 1881
Paris, Bibliothèque nationale de France, fonds français, 5036
Paris, Bibliothèque nationale de France, fonds français, 9220
Paris, Bibliothèque nationale de France, fonds français, 12483
Paris, Bibliothèque nationale de France, fonds français, 12581
Paris, Bibliothèque nationale de France, fonds français, 12603
Paris, Bibliothèque nationale de France, fonds français, 12604
Paris, Bibliothèque nationale de France, fonds français, 19152
Paris, Bibliothèque nationale de France, fonds français, 24432
Paris, Bibliothèque nationale de France, fonds français, 25545
Paris, Bibliothèque nationale de France, fonds latin, 395
Paris, Bibliothèque nationale de France, nouvelles acquisitions françaises, 10721
Paris, Bibliothèque nationale de France, nouvelles acquisitions latines, 1746
Paris, Bibliothèque Sainte-Geneviève, 2200
Valenciennes, Bibliothèque municipale, 150
Vatican city, Biblioteca Apostolica Vaticana, Ottoboniani latini, 2523

Printed works

Primary sources

Buzzetti Gallarati, Silvia (ed.). 1990. *Dit de Guillaume d'Engleterre* (Alessandria: Edizioni dell'Orso)
Bastin, Julia, and Edmond Faral (eds.). 1959–1960. *Œuvres complètes de Rutebeuf*, 2 vols (Paris: Picard)
Beltran, Evencio (ed.). 1986. *Archiloge Sophie et Livre des bonnes mœurs* (Paris: Champion)
Blangez, Gérard (ed.). 1979–1986. *Ci nous dit, recueil d'exemples moraux*, 2 vols (Paris: Société des anciens textes français)
Breul, Karl. 1895. '*Le dit de Robert le Diable*', in *Abhandlungen Herrn Prof. Dr. Adolf Tobler zur Feier seiner fünfundzwanzigjährigen Thätigkeit als ordentlicher Professor an der Universität Berlin* (Halle: Niemeyer), pp. 464–509
Brisson, Marie (ed.). 1974. *A Critical Edition and Study of Frère Robert (Chartreux) 'Le Chastel perilleux'*, 2 vols (Salzburg: Universität Salzburg)
Chase, Wayland Johnson (ed.). 1922. *The Distichs of Cato: a Famous Medieval Textbook* (Madison, WI: University of Wisconsin)
Chaurand, Jacques (ed.). 1971. *Fou, dixième conte de la 'Vie des Pères', conte pieux du XIII^e siècle* (Geneva: Droz)

Cserepfalvi, Katalin Yvonne Galligan (ed.). 1968. '*L'Arbre d'Amour*: A Critical Edition of a Previously Unedited XIV-Century Manuscript', unpublished PhD thesis (Chapel Hill: University of North Carolina)
Denomy, Alex J. 1956. 'An Old French version of the Julian episode in the *Life of Saint Basil*', *Mediæval Studies*, 18: 105–124
Dufournet, Jean (ed.). 1998. *Fabliaux du Moyen Âge* (Paris: Flammarion)
Fant, Carl (ed.). 1887. *Légende de saint Grégoire, rédaction du XIVe siècle* (Upsala: Almqvist & Wiksell)
Flutre, Louis-Fernand. 1966. '*De la pécheresse qui étrangla trois enfants*: Poème inédit de la première moitié du XIVe siècle', in *Mélanges offerts à René Crozet à l'occasion de son soixante-dixième anniversaire*, 2 vols, ed. by Pierre Gallais and Yves-Jean Riou (Poitiers: Société d'études médiévales), ii, pp. 1293–1302
Fuhrken, George E. (ed.). 1895. *De David li prophecie. Ein altfranzösisches Gedicht aus dem XII. Jahrhundert. Zum ersten Mal nach der Hs. herausgegeben* (Halle: Karras)
Gatto-Pyko, Danièle, and Harry S. Williams. 1976. '*Le Pas Salhadin*: texte et glossaire', in *Oelschläger Festschrift* (Chapel Hill: Estudios de Hispanófila), pp. 191–202
Gaucher, Élisabeth (ed.). 2006. *Robert le Diable. Edition bilingue.* (Paris: Champion)
Glixelli, Stefan (ed.). 1914. *Les Cinq Poèmes des trois morts et des trois vifs* (Paris: Champion)
Graf, Arturo (ed.). 1882–1883. *Roma nelle memoria e nelle immaginazioni del medio evo*, 2 vols (Turin: Loescher)
van Hamel, Anton Gerard (ed.). 1885. *Li Romans de Carité et Miserere du Renclus de Moiliens, poèmes de la fin du XIIe siècle* (Paris: Vieweg)
Hécart, Gabriel Antoine Joseph (ed.). 1833. *Serventois et sottes chansons couronnés à Valenciennes*, 3rd edn (Valenciennes: Prignet)
van Hemelryck, Tania (ed.). 1995. *Le Miroir de mort* (Louvain-la-Neuve: Université catholique de Louvain)
Hoepffner, Ernst (ed.). 1910. *La Prise amoureuse von Jehan Acart de Hesdin: allegorische Dichtung aus dem XIV. Jahrhundert* (Halle: Niemeyer)
Jordan, Leo. 1931. '*Le Dit du cordouanier*: notes et corrections', *The Romanic Review*, 22.4: 320–322
Jubinal, Achille (ed). 1835. *Jongleurs et trouvères, ou, Choix de saluts, épîtres, rêveries et autres pièces légères des XIIIe et XIVe siècles* (Paris: Merklein)
——— (ed.). 1839–1842. *Nouveau recueil de contes, dits, fabliaux et autres pièces inédites des XIIIe, XIVe et XVe siècles*, 2 vols (Paris: Pannier)
———. 1846. *Lettres à M. Le Comte de Salvandy sur quelques-uns des manuscrits de la Bibliothèque royale de La Haye* (Paris: Didron)
——— (ed.). 1874–1875. *Œuvres complètes de Rutebeuf, trouvère du XIIIe siècle*, 3 vols (Paris: Daffis)
Kastner, Léon-Émile. 1906. 'Les versions françaises inédites de *La Descente de saint Paul en enfer*', *Revue des langues romanes*, 49: 427–450
Koenig, V. Frederic (ed.). 1955–1970. *Les Miracles de Nostre Dame par Gautier de Coinci*, 4 vols (Geneva: Droz), ii
Kosta-Théfaine, Jean-François (ed.). 2010. *Proverbes moraulx, manuscrit Londres, British library, Harley 4431* (Clermont-Ferrand: Paleo)
Långfors, Arthur (ed.). 1906. *Li Ave Maria en roumans par Huon le Roi de Cambrai* (Helsingfors: Société néo-philologique)
———. 1917. *Les Incipit des poèmes français antérieurs au XVIe siècle* (Paris: Champion)
———. 1918. '*Le Dit de Dame Jouenne*: version inédite du fabliau du *Pré tondu*', *Romania*, 45: 99–107
———. 1921. '*Le Miroir de vie et de mort*, par Robert de l'Omme (1266), modèle d'une moralité wallonne du XVe siècle (premier article)', *Romania*, 47: 511–531
———. 1927. 'Trois *Dits de Notre Dame*: tirés du manuscrit français 24432 de la Bibliothèque nationale', *Neuphilologische Mitteilungen*, 28.4: 232–237
Lecoy, Félix. 1946. '*Le Dit de l'espervier*', *Romania*, 69: 528–534
Mareuse, Edgar (ed.). 1875. *Le Dit des rues de Paris (1300) par Guillot (de Paris)* (Paris: Librairie générale)
Marot, Clément. 1533. *Adolescence Clémentine* (Paris: Roffet)
Méon, Dominique Martin (ed.). 1823. *Nouveau recueil de fabliaux et contes inédits, des poètes français des XIIe, XIIIe, XIVe et XVe siècles*, 2 vols (Paris: Chasseriau), ii

Mercier, Alfred, and Antoine Thomas. 1894. '*La Chasse aux médisants*, poème français allégorique de Raimon Vidal', *Annales du Midi*, 6: 465–494
Meyer, Paul. 1871. 'Notice du ms. F 149 de la Bibliothèque nationale de Madrid', *Bulletin de la Société des anciens textes français*, 4: 38–59
———. 1877. 'Notice sur un ms. bourguignon (Musée britannique Addit. 15606) suivie de pièces inédites', *Romania*, 6: 1–46
———. 1912. '*Le Dit du hardi cheval*', *Romania*, 41: 90–94
Micha, Alexandre (ed.). 1994. *Marie de France: Lais* (Paris: Flammarion)
Michel, Francisque (ed.). 1836–1840. *Chroniques anglo-normandes. Recueil d'extraits et d'écrits relatifs à l'histoire de Normandie et d'Angleterre pendant les XI^e^ et XII^e^ siècles*, 3 vols (Rouen: Frère), iii
Migne, Jacques-Paul (ed.). 1844–1865. *Patrologia Latina*, 221 vols (Paris: Garnier), xi
Morawski, Joseph (ed.). 1923. *Le Facet en françoys* (Poznań: Poznańskie towarzystwo przyjaciół nauk)
——— (ed.). 1924. *Les Diz et proverbes des sages (Proverbes as philosophes)* (Paris: Presses universitaires de France)
——— (ed.). 1925. *Proverbes français antérieurs au XV^e^ siècle* (Paris: Champion)
Munk Olsen, Birger (ed.). 1978. *Dits en quatrains d'alexandrins monorimes de Jehan de Saint-Quentin* (Paris: Société des anciens textes français)
Neufang, Alfred. 1914. *Mitteilungen aus der Alexandrinerversion der Chanson d'Ami et d'Amille* (Griefswald: Adler)
Nodier, Charles (ed.). 1843. *Nouvelle bibliothèque bleue, ou légendes populaires de la France* (Paris: Belin-Leprieur)
Noomen, Willem, and Nico van den Boogaard (eds.). 1983–1998. *Nouveau Recueil complet des Fabliaux*, 10 vols (Assen: van Gorcum)
Orbán, Árpád (ed.). 1978. *Liber Floretus* (Kastellaun/Hunsrück: Henn)
Paris, Gaston, and Léopold Pannier (eds.). 1872. *La Vie de saint Alexis, poème du XI^e^ siècle et renouvellements des XII^e^, XIII^e^ et XIV^e^ siècles* (Paris: Bibliothèque de l'École des hautes études)
Paris, Gaston, and Ulysse Robert (eds.). 1876–1893. *Miracles de Nostre Dame par personnages*, 8 vols (Paris: Didot), i
de Paris, Guillot. 1866. *Les Rues de Paris mises en vers à la fin du 13^e^ siècle* (Paris: Baillieu)
Pichon, Jérôme (ed.). 1846. *Le Ménagier de Paris: traité de morale et d'économie domestique* (Paris: Crapelet)
Robert, A. C. M. (ed.). 1825. *Fables inédites des XII^e^, XIII^e^ et XIV^e^ siècles, et fables de la Fontaine rapprochées de celles de tous les auteurs qui avoient, avant lui, traité les mêmes sujets*, 2 vols (Paris: Cabin)
Scheler, Auguste (ed.). 1866–1867. *Dits et contes de Baudouin de Condé et de son fils Jean*, 3 vols (Brussels: Devaux)
——— (ed.). 1868. *Dits de Watriquet de Couvin* (Brussels: Devaux)
Schutz, Alfred H., and Nikolaus Fischer. 1931. 'An anti-feminist treatise of fifteenth-century Spain: Lucena's *Repetición de amores*', *The Romanic Review*, 22.2: 132–136
Short, Ian (ed.). 2021. *Crestien's Guillaume d'Angleterre – William of England: an edition and annotated translation* (Exeter: University of Exeter Press)
Tarbé, Prosper (ed.). 1851. *Le Tornoiement de l'Antéchrist par Huon de Méry* (Reims: Regnier)
Thornton, Hermann H. (ed.). 1932. *La Vie de saint Jehan Bouche d'or et la Vie de sainte Dieudonnee, sa mere* (Baltimore, MD: Johns Hopkins Press)
Todd, Henry A. (ed.) 1883. *Le Dit de la panthère d'amours par Nicole de Margival: poème du XIII^e^ siècle* (Paris: Didot)
Toulmin Smith, Lucy, and Paul Meyer (eds.). 1889. *Les Contes moralisés de Nicole Bozon, frère mineur* (Paris: Didot)
Trebutien, Guillaume-Stanislas (ed.). 1835. *Le Dit de ménage, pièce en vers du XIV^e^ siècle* (Paris: Silvestre)
——— (ed.). 1836. *Le Pas Salhadin, pièce historique en vers relatives aux Croisades* (Paris: Silvestre)
——— (ed.). 1837. *Le Dit des trois pommes: légende en vers du XIV^e^ siècle* (Paris: Silvestre)
Türk, Monika (ed.). 2000. *Lucidaire de grant sapientie: Untersuchung und Edition der altfranzösischen Übersetzung 1 des 'Elucidarium' von Honorius Augustodunensis* (Tübingen: Niemeyer)
Ulrich, J. 1904. 'Der Cato Jean Lefevre's nach der Turiner Handschrift I. III. 14 zum erstenmal herausgegeben', *Romanische Forschungen*, 15: 70–106
Viollet, Paul (ed.). 1881–1886. *Les Établissements de saint Louis*, 5 vols (Paris: Renouard), ii

Wallensköld, Axel (ed.). 1907–1909. *Florence de Rome, chanson d'aventure du premier quart du XIII^e siècle*, 2 vols (Paris: Didot)
Warnke, Karl (ed.). 1900. *Die Lais der Marie de France* (Halle: Niemeyer)
Weiske, Brigitte (ed.). 1991–1992. *Gesta Romanorum*, 2 vols (Tübingen: Niemeyer)
Wilmotte, Maurice (ed.). 1886. *L'Enseignement de la Philologie Romane à Paris et en Allemagne (1883–1885)* (Brussels: Polleunis, Ceuterick & Lefébure)
Wolfgang, Lenora D. 1990. *Le Lai de l'oiselet: An Old French Poem of the Thirteenth Century* (Philadelphia: American Philosophical Society)
Wolter, Eugen (ed.). 1879. *Der Judenknabe. 5 Griechische, 14 Lateinische und 8 Französische Texte* (Halle: Niemeyer)
Wright, Thomas (ed.). 1842. *A Selection of Latin Stories, from Manuscripts of the Thirteenth and Fourteenth Centuries* (London: Percy Society)
Zetterberg, Anders, and Sven Sandqvist (eds.). 1994. *Les Propriétés des choses selon le 'Rosarius' (B. N. f. fr. 12483)* (Lund: Lund University Press)
Zink, Michel (ed.). 1989–1990. *Œuvres complètes de Rutebeuf, texte établi, traduit, annoté et présenté avec variantes*, 2 vols (Paris: Garnier)

Secondary sources

Abiker, Séverine. 2008. 'L'écho paradoxal: étude stylistique de la répétition dans les récits brefs en vers, XII^e-XIV^e siècles' (Poitiers: Université de Poitiers)
Azzam, Wagih, Olivier Collet, and Yasmina Foehr-Janssens. 2010. 'Mise en recueil et fonctionnalités de l'écrit', in *Le Recueil au Moyen Âge. Le Moyen Âge central*, ed. by Olivier Collet and Yasmina Foehr-Janssens (Turnhout: Brepols), pp. 11–34
Beaussart, François-Jérôme. 2000. 'Héroïsme ou sainteté: les versions religieuses et profanes de l'histoire de Florence de Rome', *PRIS-MA*, XVI.31: 3–30
Birrer, Larissa. 2013. 'L'Exposition des songes selon Daniel', unpublished PhD thesis (Zurich: University of Zurich)
Bozzolo, Carla, Dominique Coq, Denis Muzerelle, and Ezio Ornato. 1990. 'Les abréviations dans les livres liturgiques du XV^e siècle: pratique et théorie', in *Actas del VIII Coloquio Internacional de Paleografía Latina* (Madrid: Manuel C. Díaz y Díaz), pp. 17–27
Brémond, Claude, Jacques Le Goff, and Jean-Claude Schmitt. 1982. *L'Exemplum* (Turnhout: Brepols)
Bruckner, Matilda Tomaryn. 2005. 'Conteur oral/recueil écrit: Marie de France et la clôture des *Lais*', *Op. Cit.: Revue de littératures française et comparée*, 5: 5–13
Brunel, Clovis. 1926. *La Fille du comte de Pontieu, nouvelle du XIII^e siècle* (Paris: Champion)
Bruys, François. 1751. *Mémoires historiques, critiques, et littéraires*, 2 vols (Paris: Hérissant), i
Busby, Keith. 1999. 'Fabliaux and the new codicology', in *The World and its Rival: Essays on Literary Imagination in Honor of Per Nykrog*, ed. by Kathryn Karczewska and Tom Conley (Amsterdam: Rodopi), pp. 137–160
———. 2002. *Codex and Context: Reading Old French Verse Narrative in Manuscript*, 2 vols (Amsterdam: Rodopi)
———. 2011. 'The Manuscripts of Marie de France', in *A Companion to Marie de France*, ed. by Logan E. Whalen (Leiden: Brill), pp. 303–317
Busonero, Paola. 2022. 'Quiring in manuscripts of the late Middle Ages', in *Trends in Statistical Codicology*, ed. by Marilena Maniaci, trans. by Mark Livesey (Berlin: De Gruyter), pp. 205–307
Campbell, Ian. 2007. 'Chi-squared and Fisher–Irwin tests of two-by-two tables with small sample recommendations', *Statistics in Medicine*, 26.19: 3661–3375
Cappelli, Adriano. 1982. *The Elements of Abbreviation in Medieval Latin Paleography*, trans. by David C. Heimann and Richard Kay (Lawrence, Kansas: University of Kansas Libraries)
Cerquiglini, Bernard. 1981. *La Parole médiévale* (Paris: Editions de Minuit)
———. 1989. *Éloge de la variante. Histoire critique de la philologie* (Paris: Seuil)
Cerquiglini-Toulet, Jacqueline. 2022. 'L'invention du *dit*. Quand le dit devient un écrit', in *Le Dit du berceau au tombeau (XIII^e-XV^e siècle)*, ed. by Isabelle Delage-Béland and Anne Salamon (Paris: Classiques Garnier), pp. 129–146

Cholette, Gabriel. 2022. '*Par cest essample doit entendre...* Le statut du récit dans la définition médiévale du *dit*', in *Le Dit du berceau au tombeau (XIII^e-XV^e siècle)*, ed. by Isabelle Delage-Béland and Anne Salamon (Paris: Classiques Garnier), pp. 91–108
Corbellari, Alain. 2004. 'Un problème de paternité: le cas d'Henri d'Andeli. I. Arguments littéraires', *Revue de linguistique romane*, 68: 47–56
Dahan, Gilbert. 1980. 'Les Juifs dans les *Miracles* de Gautier de Coincy (I)', *Archives Juives*, 3: 41–49, 59–68
Daly, E. J. 1950. 'An early ninth-century manuscript of Boethius', *Scriptorium*, 4.2: 205–219
Delage-Béland, Isabelle, and Anne Salamon (eds.). 2022. *Le Dit du berceau au tombeau (XIII^e-XV^e siècle)* (Paris: Classiques Garnier)
Derolez, Albert. 2003. *The Palaeography of Gothic Manuscript Books: From the Twelfth to the Early Sixteenth Century* (Cambridge: Cambridge University Press)
Dows-Miller, Sebastian. 2022. 'A new approach to scribal abbreviation in the *Bestiaire* in Merton College Library, MS 249', *Reinardus*, 34: 60–76
Driscoll, Matthew James. 2006. 'Levels of transcription', in *Electronic Textual Editing*, ed. by Lou Burnard, Katherine O'Brien O'Keeffe, and John Unsworth (New York: Modern Language Association), pp. 254–261
Drucker, Johana. 2011. 'Humanities approaches to graphical display', *Digital Humanities Quarterly*, 5.1
Ducrot, Oswald. 1984. *Le Dire et le dit* (Paris: Éditions de Minuit)
———. 1989. *Logique, structure et énonciation* (Paris: Éditions de Minuit)
Eder, Maciej, Jan Rybicki, and Mike Kestemont. 2016. 'Stylometry with R: a package for computational text analysis', *R Journal*, 8.1: 107–121
Faral, Edmond. 1910. *Les Jongleurs en France au Moyen Âge* (Paris: Champion)
Foulet, Alfred, and Mary Blakely Speer. 1979. *On Editing Medieval French Texts* (Lawrence, Kansas: Regents Press of Kansas)
Franklin, Alfred. 1867. *Les Anciennes Bibliothèques de Paris: églises, monastères, collèges, etc.*, 3 vols (Paris: Imprimerie Impériale), i
———. 1874. *Les Rues et les cris de Paris au XIII^e siècle* (Paris: Willem/Daffis)
Gossen, Charles Théodore. 1976. *Grammaire de l'ancien picard* (Paris: Klinksieck)
Gowans, Linda. 1990. '*Guillaume d'Angleterre*: prologue and authorship', *French Studies Bulletin*, 10.35: 1–5
Grente, Georges, Robert Bossuat, Louis Pichard, Guy Raynaud de Lage, Geneviève Hasenohr, and Michel Zink. 1994. *Dictionnaire des lettres françaises: Le Moyen Âge*, 2nd edn (Paris: Fayard)
Gröber, Gustav. 1888–1902. *Grundriss der romanischen Philologie*, 2 vols (Strasburg: Trübner), ii
de Hamel, Christopher. 2018. *Making Medieval Manuscripts* (Oxford: Bodleian Library)
Hasenohr, Geneviève. 1998. 'Abréviations et frontières de mots', *Langue française*, 198: 24–29
———. 2002. 'Écrire en latin, écrire en roman: réflexions sur la pratique des abréviations dans les manuscrits français des XII^e et XIII^e siècles', in *Langues et peuples d'Europe. Cristallisation des identités romanes et germaniques (VII^e-XI^e siècles)*, ed. by Michel Baniard (Toulouse: CNRS), pp. 79–110
Honkapohja, Alpo. 2021. 'Digital approaches to manuscript abbreviations: Where are we at the beginning of the 2020s?', *Digital Medievalist*, 14(1).1: 1–40
Jeay, Madeleine. 2006. *Le Commerce des mots: l'usage des listes dans la littérature médiévale (XII^e–XV^e siècles)* (Geneva: Droz)
Kedar, Benjamin Z. 1973. 'Toponymic surnames as evidence of origin: some medieval views', *Viator*, 4: 123–130
Kleist, Wolfram. 1973. *Die erzählende französische Dit-literatur in "quatrains alexandrins monorimes"* (Hamburg: Buske)
Kopaczyk, Joanna. 2011. 'A V or not a V? Transcribing abbreviations in seventeen manuscripts of the "Man of Law's Tale" for a digital edition', in *Scribes, Printers, and the Accidentals of their Texts*, ed. by Jacob Thaisen and Hanna Rutkowska (Frankfurt am Main: Peter Lang), pp. 91–106
Kunstmann, Pierre. 1991. *Miracles de Notre-Dame tirés du Rosarius (Paris, ms. B.N.fr. 12483)* (Ottawa: Presses de l'Université d'Ottawa)
Lacy, Norris J. 1987. 'Fabliaux and the question of genre', *Reading Medieval Studies*, 13: 25–34

Långfors, Arthur. 1916. 'Notice du manuscrit français 12483 de la Bibliothèque nationale', *Notices et extraits des manuscrits de la Bibliothèque nationale et autres bibliothèques*, 39.2: 503–665
Le Clerc, Victor. 1856. 'Fabliaux', in *Histoire littéraire de la France* (Paris: Didot), xxiii, pp. 69–215
Lefèvre, Yves, Philippe Ménard, Robert Martin, and Marc Wilmet. 1976–1983. *Manuel du français du moyen âge*, 3 vols (Bordeaux: SOBODI), ii
Lepage, Yvan G. 2001. *Guide de l'édition de textes en ancien français* (Paris: Champion)
Lecco, Margherita. 2020. 'Il *Dit de Merlin Mellot* fra racconto tradizionale e riscrittura edificante', *Revista de Filología Románica*: 123–129
Lote, Georges. 1949–1996. *Histoire du vers français*, 9 vols (Aix-en-Provence: Presses universitaires de Provence), ii
Marchello-Nizia, Christiane. 1979. *Histoire de la langue française aux XIV^e^ et XV^e^ siècles* (Paris: Bordas)
Marchello-Nizia, Christiane, Bernard Combettes, Sophie Prévost, and Tobias Scheer (eds.). 2020. *Grande Grammaire historique du français*, 2 vols (Boston, MA: De Gruyter), i
Marnette, Sophie. 1998. *Narrateur et point de vue dans la littérature française médiévale: une approche linguistique* (Bern: Peter Lang)
———. 2005. *Speech and Thought Presentation in French* (Amsterdam: John Benjamins)
———. 2006. 'La signalisation du discours rapporté en français médiéval', *Langue française*, 149.1: 31–47
Martin, Henry. 1885–1899. *Catalogue des manuscrits de la Bibliothèque de l'Arsenal*, 3 vols (Paris: Plon), ii
Mattiacci, Angela Marie. 1996. 'Le "Bestiaire marial" tiré du "Rosarius" (Paris, ms. B.N. f. fr. 12483)', unpublished PhD thesis (Ottawa: University of Ottawa)
Mazziotta, Nicolas. 2008. 'Traiter les abréviations du français médiéval. Théorie de l'écriture et pratiques d'encodage', *Corpus*, 7
Minnis, Alastair J. 1988. *Medieval Theory of Authorship: Scholastic Literary Attitudes in the Later Middle Ages* (Aldershot: Scolar Press)
Morawski, Joseph. 1939. 'Mélanges de littérature pieuse. II. Les miracles en quatrains alexandrins monorimes: Jean de Saint-Quentin', *Romania*, 65: 327–358
———. 1940. 'Mélanges de littérature pieuse. II. Les miracles en quatrains alexandrins monorimes (suite et fin)', *Romania*, 66: 505–129
Munk Olsen, Birger. 1963a. 'Étude et édition des *Dits* en quatrains alexandrins monorimes du manuscrit français 24432 de la Bibliothèque nationale', unpublished PhD thesis (Paris: École pratique des hautes études)
———. 1963b. 'Problèmes d'attribution: la version de *La Vie de saint Alexis* en quatrains monorimes', in *Études romanes dédiées à Andreas Blinkenberg à l'occasion de son soixante-dixième anniversaire*, ed. by Karl D. Uitti (Copenhagen: Munksgaard), pp. 144–165
Naetebus, Gotthold. 1891. *Die nicht-lyrischen Strophenformen des Altfranzösischen* (Leipzig: Hirzel)
Nixon, Terry. 1993. 'Catalogue of manuscripts', in *The Manuscripts of Chrétien de Troyes*, ed. by Keith Busby, Terry Nixon, Alison Stones, and Lori Walters, 2 vols (Amsterdam: Rodopi), ii, pp. 1–85
Nykrog, Per. 1957. *Les Fabliaux: étude d'histoire littéraire et de stylistique médiévale* (Copenhagen: Munksgaard)
Papillon, Philibert. 1742. *Bibliothèque des auteurs de Bourgogne*, 2 vols (Dijon: Marteret), i
Parussa, Gabriella, and Richard Trachsler. 2006. 'La scripta de Jacotin du Bois, un copiste dans l'atelier de Jehan Wauquelin', in *Jean Wauquelin: De Mons à la cour de Bourgogne*, ed. by Marie-Claude de Crécy (Turnhout: Brepols), pp. 185–200
Pearson, Karl. 1895. 'Notes on regression and inheritance in the case of two parents', *Proceedings of the Royal Society of London*, 58: 240–242
Pope, Mildred Katharine. 1952. *From Latin to Modern French with Especial Consideration of Anglo-Norman: Phonology and Morphology*, 2nd edn (Manchester: Publications of the University of Manchester)
Posmyk, Olaf. 2013. 'Le recueil Paris, BnF, fr. 24432 et ses copistes' (presented at the CILPR XXVII, Nancy)
Prou, Maurice. 1890. *Manuel de paléographie latine et française du VI^e^ au XVII^e^ siècle* (Paris: Picard)
van Reenen, Pieter, and Lene Schøsler. 1997. 'La déclinaison en ancien et en moyen français: deux tendances contraires', in *Le Moyen français: philologie et linguistique: approches du texte et du discours.*

Actes du VIII^e Colloque international sur le moyen français, ed. by Bernard Combettes and Simone Monsonégo (Paris: Didier Érudition), pp. 595–612
Rothwell, William. 1985. '*From Latin to Modern French*: fifty years on', *Bulletin of the John Rylands University Library of Manchester*, 68.1: 179–209
Rouse, Richard, and Mary Rouse. 2000. *Manuscripts and Their Makers: Commercial Book Producers in Medieval Paris, 1200–1500*, 2 vols (Turnhout: Harvey Miller)
Roussel, Claude. 2007. '*Le Dit des annelets* de Jehan de Saint-Quentin', in *Romans d'antiquité et littérature du nord: Mélanges offerts à Aimé Petit*, ed. by Sarah Baudelle-Michels, Marie-Madeleine Castellani, Philippe Logié, and Emmanuelle Poulain-Gautret (Paris: Champion), pp. 729–739
Rubin, Miri. 1999. *Gentile Tales: The Narrative Assault on Late Medieval Jews* (New Haven, CT: Yale University Press)
Saintsbury, George. 1911. 'Marot, Clement', *Encyclopaedia Britannica* (Cambridge: Cambridge University Press), pp. 748–749
Sandqvist, Sven. 1996. *Le Bestiaire et le lapidaire du Rosarius (B.N.f.fr. 12483)* (Lund: Lund University Press)
Simpson, James R. 2019. 'Melly and Merlin: locating little voices in Paris BnF fr. 24432', in *Shaping Identity in Medieval French Literature: The Other Within*, ed. by Adrian P. Tudor and Kristin L. Burr (Gainesville: University Press of Florida), pp. 107–120
Slautina, Maria. 2012. 'The disputed authorship of a medieval text: a new solution to the attribution problem in the case of *Guillaume d'Angleterre*', *Authorship*, 2.1
Stutzmann, Dominique. 2010. 'Paléographie statistique pour décrire, identifier, dater... Normaliser pour coopérer et aller plus loin?', in *Codicology and Palaeography in the Digital Age 2*, ed. by Franz Fischer, Christiane Fritze, and Georg Vogeler (Norderstedt: BoD), pp. 247–278
———. 2014. 'Conjuguer diplomatique, paléographie et édition électronique. Les mutations du XII^e siècle et la datation des écritures par le profil scribal collectif', in *Digital Diplomatics: The Computer as a Tool for the Diplomatist?*, ed. by Antonella Ambrosio, Sébastien Barret, and Georg Vogeler (Cologne: Böhlau), pp. 271–290
Stowell, William A. 1908. *Old French Titles of Respect in Direct Address* (Baltimore, MD: Furst)
Tubach, Frederic C. 1969. *Index Exemplorum: A Handbook of Medieval Religious Tales* (Helsinki: Suomalainen Tiedeakatemia)
Tether, Leah, Laura Chuhan Campbell, and Benjamin Pohl. 2021. *The Bristol Merlin: Revealing the Secrets of a Medieval Fragment* (Leeds: Arc Humanities Press)
Wallensköld, Axel. 1907. *Le Conte de la femme chaste convoitée par son beau-frère. Étude de littérature comparée* (Helsingfors: Imprimerie de la Société de littérature finnoise)
Walton, Thomas. 1928. 'Notes sur le manuscrit 871 de la Bibliothèque municipale de Grenoble', *Romania*, 54: 465–475
Welch, Bernard Lewis. 1947. 'The generalization of "Student's" problem when several different population variances are involved', *Biometrika*, 34.1–2: 28–35
Wellesley, Mary. 2021. *Hidden Hands: The Lives of Manuscripts and Their Makers* (London: RiverRun)
Welter, Jean Thiébaut. 1927. *L'Exemplum dans la littérature religieuse et didactique du Moyen Âge* (Paris: Occitania)
Wenzel, Siegfried. 2016. *Medieval 'Artes Praedicandi': A Synthesis of Scholastic Sermon Structure* (Toronto: University of Toronto Press)
Zufferey, François. 2004. 'Un problème de paternité: le cas d'Henri d'Andeli. II. Arguments linguistiques', *Revue de linguistique romane*, 68: 57–78

Digital resources and software

Anthony, Laurence. 2023. 'AntConc' (Tokyo: Waseda University) <https://www.laurenceanthony.net/software/antconc/> [accessed 26 October 2022]
ATILF. 2020. 'Dictionnaire du Moyen Français' <http://www.atilf.fr/dmf> [accessed 27 July 2024]
Brun, Laurent. 2023. 'ARLIMA: Archives de Littérature du Moyen Âge' <https://www.arlima.net/> [accessed 19 August 2024]
Erickson, Hal, Alan Witkowski, Anna Neatrour, Randy Silverman, and Joyce Ogburn. 2011. 'RetroReveal' <https://retroreveal.com> [accessed 2 June 2022]

Grange, Huw. 2018. 'Spreading the Light: The "Elucidarium" in French and Occitan' <http://huwgrange.co.uk/elucidarium.html> [accessed 16 March 2022]

Guillot-Barbance, Céline, Serge Heiden, Alexei Lavrentev, Matthieu Decorde, Nadine Pontal, *et al.* 2022. 'Base de Français Médiéval' (Lyon: ENS de Lyon) <http://bfm.ens-lyon.fr/> [accessed 19 August 2024]

Lass, Roger, and Margaret Laing. 2013. '2. Interpreting Middle English', in *An Electronic Version of a Linguistic Atlas of Late Mediaeval English*, ed. by Michael Benskin, Margaret Laing, Vasilios Karaiskos, and Keith Williamson (Aberdeen: Aberdeen University Press) <http://www.lel.ed.ac.uk/ihd/elalme/elalme.html> [accessed 14 August 2023]

Lavrentiev, Alexei. 2023. 'Logiciel RNNTagger' <http://bfm.ens-lyon.fr/spip.php?article355> [accessed 5 October 2023]

Lexical Computing. 2023. 'Sketch Engine'. (Brno: Lexical Computing) <https://www.sketchengine.eu/> [accessed 17 August 2024]

OpenAI. 2022a. 'ChatGPT' <https://chat.openai.com/chat> [accessed 17 August 2024]

———. 2022b. 'Introducing ChatGPT' <https://openai.com/blog/chatgpt> [accessed 12 August 2023]

Piccard, Gerhard. 1961. 'Wasserzeichensammlung Piccard', *Hauptstaatsarchiv Stuttgart* <https://www.piccard-online.de/struktur.php> [accessed 6 June 2024]

Pierreville, Corinne. 2023. 'Corpus FABLIAUX-V0-1' <https://txm-bfm.huma-num.fr/txm/> [accessed 12 October 2023]

Sentjens, Cedric. 2023. 'Rutebeuf.be' <http://www.rutebeuf.be> [accessed 17 August 2024]

Schmid, Helmut. 1994. 'Probabilistic Part-of-Speech Tagging Using Decision Trees', in *Proceedings of the International Conference on New Methods in Language Processing* (Manchester) <https://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/data/tree-tagger1.pdf> [accessed 7 September 2023]

———. 2019. 'RNNTagger' (Munich: Ludwig-Maximilians-Universität) <https://www.cis.uni-muenchen.de/~schmid/tools/RNNTagger/> [accessed 6 October 2023]

Seláf, Levente. 2011. *Le Nouveau Naetebus* <https://nouveaunaetebus.elte.hu/> [accessed 26 May 2023]

Appendices

Appendix A: fr24432 contents

Item	Folios	Short Title	Full Title	Author[1]	Description
	0ra–0va				Table of contents
1	1ra–13va	*Guillaume*	*Le dit de Guillaume d'Engleterre*	Jehan de Saint-Quentin	An alexandrine reworking of the epic tale of *Guillaume d'Angleterre*.
2	13va–14rb	*Trois mors*	*Le dit des trois mors et III vis*		A discussion between three living and three dead men on the importance of living a good life.
3	14rb–17ra	*Respon*	*Un dist que on clamme respon*	Jehan de Saint-Quentin	A girl is saved from rape, and a canon from death, by their devotion to Mary.
4	17ra–22vb	*Mesdisans 1*	*La chace des mesdisans*	Raimon Vidal	The allegorical tale of a hunt for an allegorical boar, representing a 'mesdisant'. Repeated at item 86.
5	22vb–25rb	*Dorée*	*La lande dorée*	Le vicomte d'Aunoy	The author-narrator tells of discovering a beautiful lady in a forest idyll, who agrees to be his lover if he completes a certain task. The text ends here.
6	25rb–25vb	*Repentance*	*La repentance Rutebeuf*	Rutebeuf	Rutebeuf, apparently close to death, expresses regret for his life choices.
7	25vb–29vb	*Desputoisson*	*La desputoisson du vin et de l'iaue*		A debate in which wines of different types plead their case before the God of Love, before water wins the day.
8	29vb–33vb	*Salhadin*	*Le pas Salhadin*		A narrative concerning key events from the Third Crusade.
9	33vb–35va	*Erberie*	*L'erberie*	Rutebeuf	A herb seller sells his saucy wares.
10	35vb–36vb	*Charlot*	*Le dit de Charlot et du barbier*	Rutebeuf	A debate between a Christian and a Jew.
11	36vb–42ra	*Chanoinnes*	*Le dit des trois chanoinnes*	Jehan de Saint-Quentin	Three canons choose a life of penitence and help a sinning knight achieve redemption.
12	42ra–45rb	*Oiselet*	*Le lai de l'oiselet*		An anti-*vilain* tale in which a rural idyll is ruined by a peasant's wicked ways.
13	45va–46va	*Complainte*	*La complainte Rutebeuf*	Rutebeuf	A short poem by Rutebeuf in which the author laments his lot.
14	46va–47vb	*Traverces*	*Le dit des traverces*		Nonsense poetry presented in the first person, with reference to locations and religious figures.
15	47vb (fragment)	*Vilains*	*Les XXIII manieres de vilains*		A fragmentary text of which only the rubricated incipit survives. Anti-peasant satire.
16	49ra (fragment)	*Vin*	*Le Paternostre du vin*		A fragmentary text of which only the final 24 lines survive. It appears to be a Goliardic parody on the words of the *Pater Noster*, with stanzas introduced by words of the Latin.
17	49ra–49vb	*Chapel*	*Le chapel des VII flors*		A lyric poem in which advice on seven feminine virtues is given through the allegory of flowers.
18	49vb–52ra	*Boivin*	*Boivin de Provins*		The manuscript's only *fabliau*, about a jongleur who gets one over on a group of unscrupulous prostitutes.
19	52rb–53va	*Ave Maria*	*Ave Maria*		The words of the Latin Ave Maria each taken to create a stanza in praise of Mary.
20	53va–56vb	*Pommes*	*Le dit des III pommes*	Jehan de Saint-Quentin	Two friends, united by a pilgrimage miracle, both sacrifice something for the other.
21	57ra–84ra	*Aristote*	*Li livres des decrés Aristote*		Pseudo-Aristotelian prose purporting to contain advice from Aristotle to Alexander on how to be a good king.
22	84rb–86va	*Planetes 1*	*Le dit des planetes*		Social criticism through comparison of seven types of people with the seven planets and seven days of the week. Repeated at item 64.
23	86va–87ra	*Espervier*	*Le dit de l'espervier*		Invective against inconstant, avaricious women.
24	87rb–91rb	*Bible*	*La grant bible Nostre Dame*		Praise of the Virgin Mary interspersed with anti-semitic invective.
25	91va	*Dormant*	*Les VII dormant*		A prose list of 'the seven sleepers.'
26	91va–99ra	*Vision*	*La vision de saint Paul*		A vernacular, alexandrine version of the *visio sancti Pauli*.
27	99ra–99rb	*Jorz*	*Les XXX jorz perilleus de l'an*		A prose list of the 30 days of the year that represent the most peril.
28	99rb–102rb	*Romme*	*Le dit de la bourjoise de Romme*	Jehan de Saint-Quentin	A pious woman is tempted to incest and kills her child, but is saved by Mary.

[1] This column includes both those texts that contain a specific reference to an author as well as those which have been assigned authorship by later scholarship. The attribution of *Guillaume* and *Robert* to Jehan de Saint-Quentin is my own, as discussed in the analysis within Chapter 3.

Item	Folios	Short Title	Full Title	Author[1]	Description
29	102rb–104va	*Vie du monde*	*La vie du monde*	Rutebeuf	Rutebeuf offers satirical criticism of the three estates.
30	104va–107rb	*Deux chevaliers*	*Le dit des II chevaliers*	Jehan de Saint-Quentin	A tale about the fates of two knights, one who is pious and one who is not.
31	107rb–109vb	*Enfant rosti*	*Le dit de l'enfant rosti*	Jehan de Saint-Quentin	The devil roasts the child of a pious couple, but Jesus is on hand to save the day.
32	109vb–112ra	*Povre chevalier*	*Le dit du povre chevalier*	Jehan de Saint-Quentin	Mary takes the place of a woman to save her and her husband from the devil.
33	112ra–115rb	*Chevalier et escuier*	*Du chevalier et de l'escuier*	Jehan de Saint-Quentin	A knight and his squire suffer for denying their faith in return for earthly reward.
34	115rb–118rb	*Narbonne*	*Le dit de la borjoise de Narbonne*	Jehan de Saint-Quentin	A widow saves her son from execution by appealing to Mary.
35	118rb–120va	*Chevalier hermite*	*Le dit du chevalier qui devint hermite*	Jehan de Saint-Quentin	A sinful knight undergoes trials to redeem his soul before becoming a hermit.
36	120va–122vb	*Cordouanier*	*Le dit du cordouanier*	Jehan de Saint-Quentin	The devil leads a cordwainer to kill his wife, but Mary revives her and he is saved.
37	122vb–125ra	*Juitel*	*Le dit du petit Juitel*	Jehan de Saint-Quentin	A Jewish child converts to Christianity.
38	125ra–128ra	*Enfant sauva mere*	*Le dit de l'enfant qui sauva sa mere*	Jehan de Saint-Quentin	A child saves his mother from damnation by performing her penance after her death.
39	128ra–130rb	*Eaue et vergier*	*Le dit de l'eaue beneoite et du vergier*	Jehan de Saint-Quentin	A bipartite tale in which a convert discovers the true power of holy water, and then a greedy man is turned to the light by a dream.
40	130rb–132va	*Pain*	*Le dit du riche home qui geta le pain a la teste du povre*	Jehan de Saint-Quentin	An avaricious man realises the error of his ways after throwing bread at a peasant's head.
41	132va–135rb	*Mescreant*	*Le dit du chien et du mescreant*	Jehan de Saint-Quentin	A man learns the power of the Host by trying to feed it to his dog, which devours him instead.
42	135rb–137rb	*Pecherresse*	*Le dit de la pecherresse qui estrangla III enfans*	Jehan de Saint-Quentin	Mary saves a pious woman who is tempted into incest and kills her children.
43	137rb–138va	*Rebellion*	*Le dit de la rebellion d'Engleterre et de Flandres*		Anti-English invective written as advice for a king.
44	138vb–142vb	*Mais*	*Le dit des mais*	Gieffroy	A poem explaining how all roles in society are important, but greed so often gets the better of us all.
45	142vb–146va	*Bacchus*	*Le martyre de saint Bacchus*	Gieffroy	A parodic hagiography of Saint Bacchus (that is, wine).
46	146va–148rb	*Requeste*	*La requeste des freres meneurs sus le septiesme Climent le quint*		Satirical criticism of monastic orders, addressed to Pope Clement V.
47	148rb–152ra	*Patenostres*	*Le dit des patenostres*	Gieffroy	Discussion of, and prayers for, all the estates.
48	152ra–153rb	*Nostre Dame*	*Le dit de Nostre Dame*		Three separate poems, all in praise of the Virgin Mary.
49	153va–171rb	*Panthere*	*La panthere d'amors*	Nicole de Margival (the *explicit* references a Richart de Fournival)	A dream allegory on bestiary themes, in which the narrator learns the significance of the panther and other beasts, as well as flowers and stones.
50	171rb–184ra	*Isopet*	*Isopet*		A version of the *Ysopet* (Paris II) containing 40 different tales.
51	184ra–198Grb	*Tournoiement*	*Dit du tournoiement d'Antecrist et d'euls de son barnage*	Huon de Méry	An allegorical tourney between the forces of the Light and the Dark.
52	198Gva–198Grb (fragment)	*Lai d'amour 1*			An untitled *lai d'amour*, of which the end is missing.
53	199ra–201vb	*Merlin*	*Le dit de Merlin Mellot*	Jehan de Saint-Quentin	A peasant receives riches from Merlin but loses them through his lack of respect.
54	202ra–215rb	*Robert*	*Le dit de Robert le Diable*	Jehan de Saint-Quentin	An alexandrine reworking of the epic tale of Robert the Devil.
55	215rb–226va	*Florence*	*De Florence de Romme*	Jehan de Saint-Quentin	An alexandrine reworking of the epic tale of Florence of Rome.
56	226va–230vb	*Menage*	*Le dit de menage*		A debate in which three speakers discuss the merits and disadvantages of married life.
57	231ra–241rb	*Anelés*	*Le dit des anelés*	Jehan de Saint-Quentin	A wife cheats on her husband while on pilgrimage but is eventually redeemed.
58	241rb–245ra	*Yonec*	*Yonec*	Marie de France	One of Marie de France's *lais*, in which a bird-knight's tragic death at the hands of his lover's jealous husband is eventually avenged by his son.
59	245ra–246vb	*Divisions*	*Les divisions de LXXII biautés qui sont en dames*		A list of the 72 beauties to be found in women, most of which are physical.

Item	Folios	Short Title	Full Title	Author[1]	Description
60	246vb (fragment)	*Trois mortes*	*Li dis des trois mortes et des trois vives*		A fragmentary version of item 2, on this occasion with female characters, of which only 12 lines survive.
61	247ra–257va	*Buef*	*Le dit du buef*	Jehan de Saint-Quentin	Three sinners, united through incest, are saved by penance.
62	257va–261vb	*Rues*	*Le dit des rues de Paris*	Guillot de Paris	A guided tour of the streets of Paris.
63	261vb–262vb	*Lendit*	*Le dit du Lendit rimé*		Praise of merchants, in particular those of the Lendit market.
64	263ra–265rb	*Planetes 2*	*Le dit des planetes*		Another version of item 22.
65	265rb			Perrenette de Pluvot	A later addition signed in a 16th-century hand.
66	265rb–279rb	*Arbre*	*Le dit de l'arbre d'amours et de touz ses [fruits] bons et mauvais*	Raimont Badaut	An allegory in which the narrator tastes the sweet and bitter fruits of the tree of love.
67	279va			Denis Lucey	A later addition signed in a 16th-century hand.
68	279rb–280ra	*Chastoi*	*Le dit du chastoi du jone gentil homme*	Jehan de Condé	An exhortation for young men to eschew sin and choose the path of righteousness.
69	280rb–281rb	*Compaignons*	*Le dit des II loiaus compaignons*	Jehan de Condé	A tale about two imprisoned men who are freed as a result of their loyal bond.
70	281rb–281va	*Despensier*	*Le dit du vilain despensier*	Jehan de Condé	A warning against being too charitable towards the poor.
71	281vb–302vb	*Songes*	*Li livres des songes Daniel et les songes Macrobe*		Prose discussion of the meaning of dreams.
72	303ra–310vb	*Serventois*	*Serventois a Valenchiennes*		A series of short songs, often on themes of love and devotion.
73	311ra–312vb	*Lai d'amour 2*			Another untitled love poem.
74	312vb–315va	*Beguine*	*Le dit de la beguine*	Jehan de Saint-Quentin	A rich beguine is damned for a deal she makes with the Devil involving the Host.
75	315va–316vb	*Anel*	*Le dit de l'enfant qui mist l'anel que s'amie li avoit donné ou doit a l'ymage de Nostre Dame*	Gautier de Coinci	Part of Gautier de Coincy's *Miracles de Nostre Dame*, in which a young *clerc* promises himself to Mary but then forgets her, before being reminded of his promise.
76	316vb–317rb	*Cheval*	*Le dit du hardi cheval*		A parodic attempt to sell an unsaleable horse.
77	317rb–327rb	*Messe*	*Le tretié sus la messe qui est dit mirouer d'iglusse que frere Hue de l'ordre des Prescheurs ordena*		An anonymous prose translation of Hugues de Saint-Cher's *Speculum ecclesiae*.
78	328(bis)ra–385rb	*Lucidaires*	*Lucidaires*		A prose debate between a master and his student on matters of faith, with an extension to the original text beginning on f. 361ra.
79	388va–390ra	*Preu chevalier*	*Le dit du preu chevalier*	Watriquet de Couvin	An allegorical tale in which a knight visits all the values until eventually reaching prowess.
80	390ra–391va	*Iraygne*	*Le dit de l'iraygne*	Watriquet de Couvin	A dream allegory in which the narrator observes a spider fighting a toad, with Reason on hand to explain the meaning.
81	391va–392va	*Drois*	*Le dit des drois*	Clerc de Vaudoy	A collection of proverbs and sayings explaining good behaviour.
82	392va–394va	*Roy*	*Le dit du roy*	Watriquet de Couvin	Advice addressed to the King of the French, summarising the virtues that Aristotle explained to Alexander (cf. item 21).
83	394va–396va	*Nois*	*Le dit de la nois*	Watriquet de Couvin	An encomium to nuts, which are said to represent all positive virtues.
84	396va–407rb	*Prise*	*La prise amoureuse de jonesce*	Jehan Acart de Hesdin	An allegory in which the narrator is the object of a hunt by Love and his entourage.
85	407rb			Perrenette de Pluvot / Clément Marot	A later addition in a 16th-century hand.
86	407rb–412rb	*Mesdisans 2*	*Le dit de la chace des mesdissans*	Raimon Vidal	Another version of item 4.
87	412rb–414va	*Jouenne*	*Le dit de dame Jouenne*		A comic dialogue in which Jouenne's husband punishes her for her shrewishness.
88	414va–419vb	*Miroir*	*Le miroir de vie et de mort*	Robert de l'Omme	A dream allegory in which the Seven Deadly Sins discuss their power over mankind.
89	420ra–436vb	*Salomon*	*Les proverbes de Salomon*		Prose presentation, with commentary, of the words of a number of saints and prophets. There are indications of an attempted presentation as dialogue.
90	437ra–443vb	*Misere*	*La misere de l'omme*		An anonymous prose translation of *De misera humanae conditionis.*

Appendix 1A: Data

Table 1.46 – variation in column length per scribal stint.

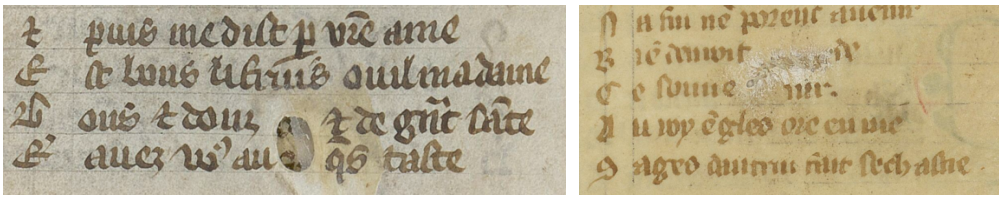
Scribe	μ (calculated)	Standard Deviation with calculated μ	μ (standardised)	Standard Deviation with standardised μ
1	38.00	0.00	38	0.00
2	38.00	0.00	38	0.00
3	38.15	0.38	38	0.41
2	38.00	0.00	38	0.00
3	38.02	0.13	38	0.13
4	37.85	0.38	38	0.41
3	37.95	0.22	38	0.23
5	38.00	0.00	38	0.00
6	38.13	1.33	38	1.34
7	38.01	0.09	38	0.09
8	38.00	0.00	38	0.00
3	37.75	0.44	38	0.50
7	38.04	0.34	38	0.34
5	38.00	0.00	38	0.00
9	38.00	0.00	38	0.00
10	36.00	0.00	36	0.00
2	38.00	0.00	38	0.00
11	38.00	0.00	38	0.00
2	38.00	0.00	38	0.00
11	37.82	1.84	38	1.84

Appendix 1B: Further codicological observations

1B.1 Damage and repairs

Unrepaired damage to the parchment dating prior to the copying of text is rare but does appear, alongside a greater number of holes resulting from the parchment preparation process that have been fixed by stitching.[1]

Figure 1.64 – left: an unrepaired pre-scribal hole on f. 274rb; right: a pre-scribal stitched repair on f. 137va.



A number of folios are missing corners, in most cases probably representing the original line of the animal skin before it was worked into parchment. This is true of the missing corner of f. 107, for example, where the parchment by the missing corner is thicker than elsewhere and darker in colour, which is consistent with it being the original hide edge.

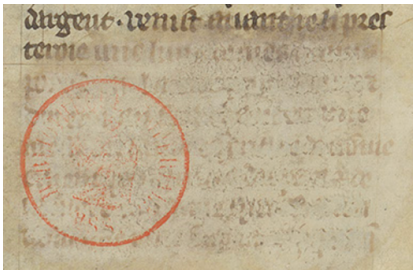
Figure 1.65 – left: the missing corner of f. 107va; right: the missing corner of f. 5rb.

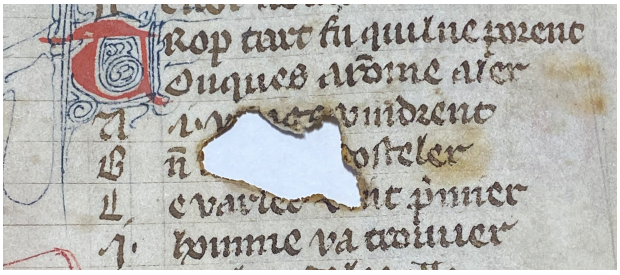
In some cases missing corners appear to be post-scribal, as with f. 5. In this case, a sewn fix on the bottom edge of the folio appears to have failed, likely sometime after production, and the corner of the folio has torn away entirely.

[1] Alongside f. 137, stitched repairs appear also on ff. 72, 113, 121, 130, 131, 147, 249, 276, 396, 430, and 431.

Evidence of other post-compilatory damage is rare in fr24432, although there are a few examples of water damage and small holes apparently made after the text was copied.

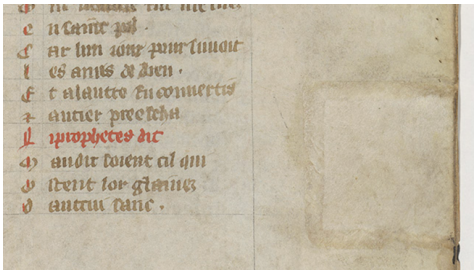
Figure 1.66 – left: water damage on f. 35rb, with a later library stamp; right: a post-scribal hole on f. 254rb.





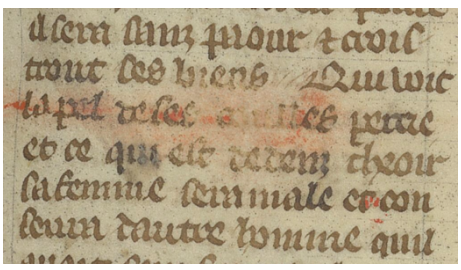
In some cases, post-scribal damage has been repaired, such as through the use of a patch on f. 426, clearly a post-scribal change given the absence of pricking marks on the patch itself.

Figure 1.67 – a post-scribal patched repair on f. 426rb.

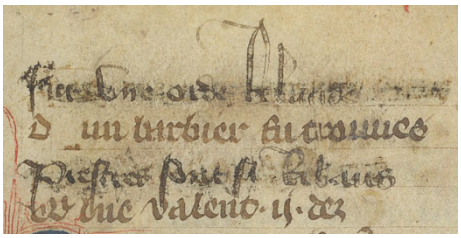


There are also some instances of later erasure of individual lines of text, likely attempts at the censure of material deemed unacceptable. This erased text is primarily limited to sexual and anti-clerical content.

Figure 1.68 – top: erasure of sexual content on f. 294ra; bottom: erasure, with later re-addition, of anti-clerical invective on f. 230ra; both with accompanying semi-diplomatic transcription.



il sera sanz paour *et* croistront ses biens [¶] Qui voit la pel de ses [cou]lles percié et ce qui est dedenz cheoir sa femme sera male et con[-]sevra d'autre homme qu'il



[Sur une orde ribaude]
D'un barbier fu trouvés
[Prestres sont si ribaus]
Qu'il ne valent .ij. dez

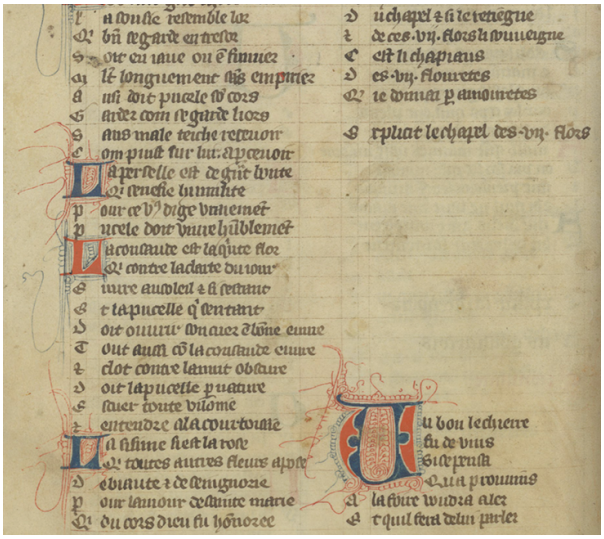
At most, then, the post-scribal damage in fr24432 indicates those areas of text which later readers found problematic, while the limited pre-scribal damage suggests that fr24432 was a manuscript in which aesthetic considerations were important, although not necessarily to the extent of needing every leaf to be free from flaws.

Appendix 1C: Further paratextual observations

1C.1 Decoration

The codex does not contain any coloured illustrations, although the original intention does appear to have been for the manuscript to be more highly decorated. Most texts are preceded by a section of blank space, usually between ten and fourteen lines in length, which appears intended as a space to represent the contents of the coming text through illustrations that were never completed.

Figure 1.69 – f. 49ᵛ, including a gap between texts possibly intended to be filled by illustration.



This hypothesis, that illustrations were planned for the codex but were never completed, is supported by the presence of references to assumed illustrations that appear within some of the texts of the manuscript, in some cases corresponding to a gap. As discussed in our examination of the *litterae notabiliores* in the section from page 40, for example, the first copy of *La chace des mesdisans* within the codex, at ff. 17ʳᵃ–22ᵛᵇ, contains rubricated captions explaining the contents of illustrations which were apparently planned but never existed (Figure 1.24).

Similarly, at the beginning of the main prologue to *Miroir*, a text examining life, death, and the seven deadly sins through an allegorical tree, the narrator references ‘Li arbres que veu avez’ (f. 414ᵛᵇ) despite no such illustration being present. This appears to correspond to an

image that may at some stage have been planned, rather than an appeal to the reader's imagination, since there is a large gap of some 14 lines left immediately before the start of the text on the same folio.

The illustration which would likely have filled this gap is found in at least two other manuscripts.[1] As is visible in Figure 1.70, there are sufficient similarities between these two illustrations, both in their imagery and the text they contain, to suggest that they are examples of a standard illustration that was in wide circulation alongside the text of *Miroir*. Indeed, the allegory may have been sufficiently well known to readers that the illustration was adequate on its own, as is the case with BnF fr. 9220, which does not contain the text to which its illustration alludes.

Figure 1.70 – examples of the tree referenced in Miroir *found outside of fr24432.*

BSG 2200, f. 164v

BnF fr. 9220, f. 10r

[1] BSG 2200 dates from the 13th century, while BnF fr. 9220 is either 13th- or 14th-century, and so both are of an age roughly contemporary with the likely exemplars which our scribes could have had access to. In a third manuscript containing the text (BnF fr. 834), this time seemingly roughly contemporary with our codex, no such illustration exists, nor does the prologue containing a reference to a tree seen by the reader. I have been unable to reliably determine the situation of a fifth manuscript (BAV Otto. lat. 2523), although this is comfortably dated to between 1450 and 1460, roughly a century after our codex was compiled, and is therefore of minimal relevance here.

It is therefore notable that this illustration is absent from fr24432, despite the presence of an appropriate gap before the start of its copy of *Miroir*, in that it suggests that the contents of the manuscript may have been planned based on an assumption that illustrations would be present, only for this to prove unfeasible or otherwise undesirable at some point in the process of production. Since *Miroir* appears in the same quire as the second copy of *La chace des mesdisans*, which contains fewer gaps for planned illustrations than the copy found towards the beginning of the codex, this supports a hypothesis that the scribe behind its completion (hand 2) had become aware by this point of a decision not to illuminate the manuscript, in contrast to what may have originally been intended when its contents were planned.

Despite the apparent lack of the originally planned illustrations, the manuscript is not without visual interest. A number of doodles by the manuscript's scribes offer an insight into the workings of the *scriptorium* in which the codex was produced.

Figure 1.71 – left: a face incorporated into catchwords on f. 8v; right: a face within the first line of f. 9ra.

On folios 8v and 9r, for example, a face appears in the embellished initial of both the catchwords at the end of q. 1, and the first line of the following quire (q. 2) to which the catchwords correspond, in both cases seemingly by hand 2 (Figure 1.71). The presence of an embellished initial at the beginning of a quire, while rare in the codex as a whole, does point towards the possible importance of the separation of quires, to hand 2 at least, as marking the end of one period of work and the beginning of another.

That the first face is seemingly more detailed than the later one is also notable, as it raises the possibility that the face at the beginning of q. 2 was a doodle as hand 2 began working, but

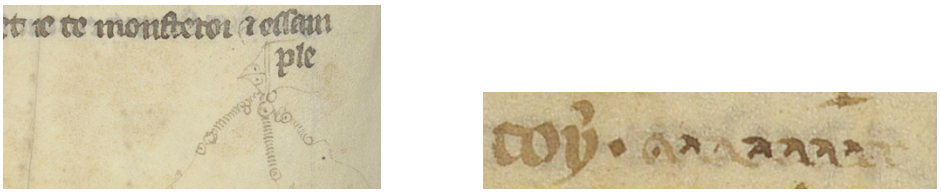
when later filling in the catchwords for the completed stint, the scribe spent a little longer improving upon their initial design. This therefore suggests that the inclusion of catchwords was a task that the scribe carried out at the end of a period of work.

A penchant for embellishment by hand 2 is also visible in the presence of another character containing a face on f. 6ra, as well as two embellished frames around the catchwords between quires 2 and 3, and 54 and 55 (Figure 1.72).

Figure 1.72 – left: an embellished m *on f. 6ra; centre: embellished catchwords on f. 16v; right: faces in embellished catchwords on f. 411v.*

Hand 2 is not the only scribe to display some artistic flair, with other scribes using doodles as a means of consolidating certain elements of the text block. For example, hand 3 embellishes and therefore brings attention to a hanging indent on f. 75vb, while a somewhat more understated geometric doodle is used by hand 5 to fill blank space in a line (Figure 1.73).

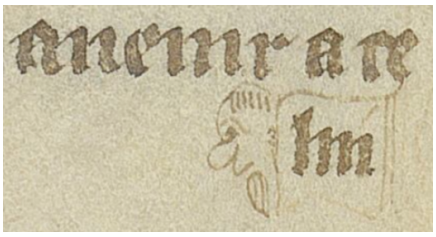
Figure 1.73 – doodles on f. 75vb (left) and 326vb (right).



Amongst those illustrations contemporary with the production of the codex, perhaps of greatest interest is the face of a tonsured man, decorating a hanging indent on f. 60rb (Figure 1.74), within the first stint of hand 3. While there are no clear conclusions to be drawn from this figure, it is nevertheless the closest we get to evidence about the identities of those responsible for the manuscript's creation. It is unclear whether the man depicted is a monk or a cleric, and whether he is a figment of the illustrator's imagination or the scribe working in

the next stall. Nevertheless, it does provide us with an image of the illustrator, most likely also the scribe of this section, as perhaps a man from a clerical background, working alongside others, which fits with the interpretation given by the evidence we have seen elsewhere that the manuscript was produced by a relatively large, organised *scriptorium*.

Figure 1.74 – a tonsured man on f. 60ᵛᵇ.



Those contemporary with the manuscript's production are not the only individuals to have contributed to its decoration, however, and a faint motif of a crown atop an 'M,' found at the bottom of f. 67ʳ in the midst of *Aristote*, a vernacular translation of the *Secreta Secretorum* (Figure 1.75), is of particular note.[2] The sketch is drawn in lead point, but not the same as was used to complete the ruling on some of the manuscript's quires, and thus this addition appears later than other illustrations elsewhere in the codex.

Figure 1.75 – left: a crown motif sketched at the bottom of f. 67ʳ; right: spectral modification of the same image using retroReveal *(Erickson* et al.*, 2011) to better reveal the pattern.*

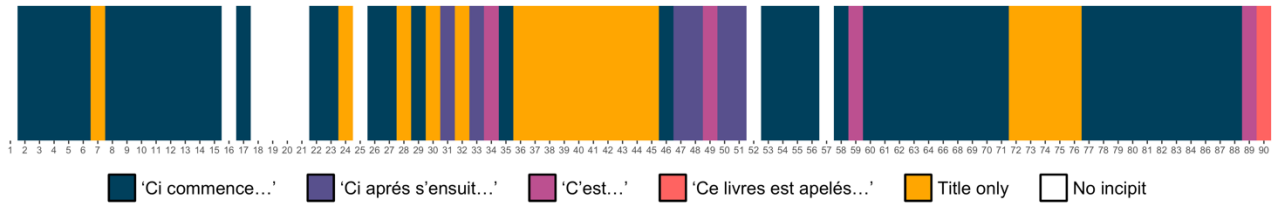
[2] It has not been possible to determine the significance of the 'M,' which is likely an initial, with the wider motif functioning as a heraldic device, given that it appears not to align with any of the individuals known to have owned fr24432 (see p. 59). Similar motifs appear in the collection of watermarks compiled by Gerhard Piccard (1961, items 28910–28918), but given that a crown atop a letter is not a particularly complex image, this is insufficient evidence to suggest that they share a common origin with the illustration here.

The presence of a crown could be seen to reveal the interests and allegiances of later readers of the codex, since the primary apparent function of the *Aristote*, as a pseudo-Aristotelian text in which the philosopher passes on wisdom to his student Alexander, is to offer advice to rulers about how best to behave and manage their realms. There is no indication that the readers of the codex (intended or otherwise) were of such high status, indeed the absence of the planned illustration would refute such a claim, but it is certainly possible that a later reader was inspired by the regal content of the text and felt that their sketch was appropriate.

1C.2 Rubrication

Rubrication is used sparingly in the body of fr24432, and primarily appears as the *incipit* marker at the beginning of a text, with rubrication therefore playing a role in the navigation of the manuscript. In most cases, the rubrication is completed in the same hand as the scribe for any given section, but the method of rubrication is generally not consistent within the work of a single scribe and instead varies, likely depending on the usage found in the exemplars from which the scribes were copying.

Figure 1.76 – distribution of different incipit *types within fr24432.*[3]



The visualised data in Figure 1.76 reveal the five different methods of marking the beginning of a text which are found across the manuscript. Most frequent is the method involving some variant of 'Ci commence' followed by the text's given title, although for *Chapel* (item 17) we are only given 'Ci comance' and the title is left out, suggesting that the scribe for that

[3] I have excluded the later additions at items 65, 67, and 85 from this visualisation. Notably, the incipit lines for item 72, as well as internal titles for the individual short poems it contains, are written in brown ink by the same hand who copied the text itself.

particular section may have added the *incipit* at the end of a prior period of work, meaning to come back and complete it once the next text had been copied. An alternative method is seen in the use of phrases akin to 'Ci aprés s'ensuit,' although this method is much less frequent, while other minor methods include the use simply of 'C'est,' or in the case of *Misere* (item 90), 'Ce livres est apelés.'

The second most frequent method of marking the beginning of a text using rubrication, however, is simply to state the text's title, and it is this that reveals the possibility that these *incipit* lines were determined by the exemplar manuscripts from which the texts were drawn. The consistent use of the title-only method for items 36–45, for example, adds another similarity between these texts, already united by their comparable content and form,[4] and suggests that they may have been copied from the same exemplar manuscript. We may also suggest that other groupings, such as items 16–21, 47–51, and 72–76, may have been copied in continuous or near-continuous runs from individual exemplars.

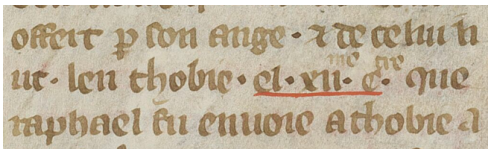
The position of rubrication when used to mark the beginnings of texts is also of interest. In most cases, the *incipit* lines are placed as early as possible, with a gap left before the text begins, possibly for illustration that was never completed. There is a good deal of variation in the length of this gap, with features such as column and page breaks having an impact on what the scribes may have considered aesthetically logical, although it is usually somewhere around twelve lines. There are some texts, however, which contain such a gap before the text begins, with the rubrication placed at the end of the gap rather than at the beginning.[5] Since all these texts are later on in the manuscript, it could be possible that for those which fall within the quires of the original base manuscript, this offers evidence of a growing realisation by the scribes involved in its production that it was most likely not going to contain

[4] On which, see the section on versification and rhyme from p. 306 below.
[5] Items 53, 54, 61, 64, 66, 69, 70, 78, 89, and 90.

illustrations. It is of course equally possible that this results from the trends present in the scribes' exemplar manuscript(s).

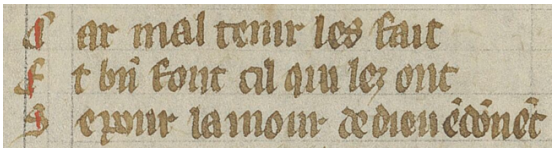
Rubrication beyond the *incipit* lines is rare but does occur. The table of contents, for example, uses rubrication for folio references to distinguish them from the titles of individual texts, while *Dormant* (item 25), a short prose text of only eight orthographic lines, is written entirely in red ink. More predictable uses of rubrication appear primarily in the manuscript's longer prose texts, but also in a number of long verse texts, too. The use of rubrication to indicate captions or other headings is observed in items 4, 21, 48, 50, 65, 71, 77, 78, 84, 86, and 90, while items 2, 78, 88, 89, and 90 also use red ink to mark a change in speaker within the texts' dialogic structures. Quotations in Latin and references to scripture often appear in red in *Messe* and *Lucidaires* (items 77 and 78), although in some cases in *Messe* this highlighting is conducted through red underlining, rather than the copying of full text in red.

Figure 1.77 – an example of a scriptural reference underlined in red, on f. 324ra.



Salomon and *Misere* (items 89 and 90), both prose texts, are interesting in that most majuscule letters they contain are highlighted using flecks of red. This method of highlighting the beginning of new clauses is used consistently within quires 56 and 58, but not in quire 57, which suggests that this particular element of the rubrication of these sections of the manuscript was not completed at the same time as the text was copied, and also that the quires were seen as a discrete unit to be attended to during separate periods of work.

Figure 1.78 – an example of red flecks of ink on capital letters, on f. 421va.



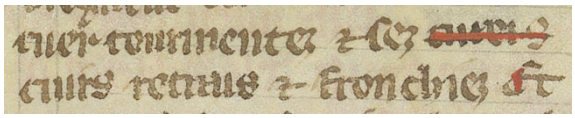
The fact that other headings and labels in red do still appear in q. 57, apparently in the hand of the scribe who completed that section, suggests perhaps that more than one person was involved in the rubrication of this section, with the main scribe completing their rubrication before passing the quire on to a less experienced scribe to highlight the capital letters. It is unlikely that even the main scribes for a given section completed their rubrication immediately, however, since *Salomon* in particular contains a number of notes as to what should be included in the rubrication, suggesting that when returning to complete it, the scribe no longer had access to the exemplar manuscript from which the rest of the text had been copied.

The scribes also appear to have used the time spent adding rubrication as an opportunity to proofread the text. There are four separate occasions where red ink is used to cross out a word that forms part of the main text, all but one of which appear within *Misere*.[6] In the earliest example the use of red would make sense, given that the rubricator is removing text that should seemingly have been copied in red, whereas in the later cases there is no apparent need for red to be used. In all cases these appear to be errors that the scribe noticed while they copied the text, with the correct form offered immediately after, and so the act of crossing out in red does appear to be a conscious choice, with hand 11 choosing highlight the correction of their earlier error using red.

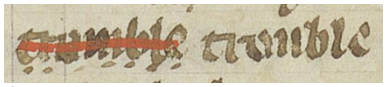
[6] ff. 325ra, 440va, 441va, 441vb. Red ink is also used to cross out text on f. 435ra, but in that case it is to correct an error in the rubrication itself.

Figure 1.79 – errors struck through using red ink.

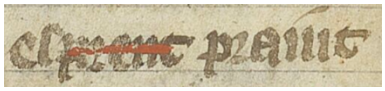
f. 325ra



f. 440va



f. 441va



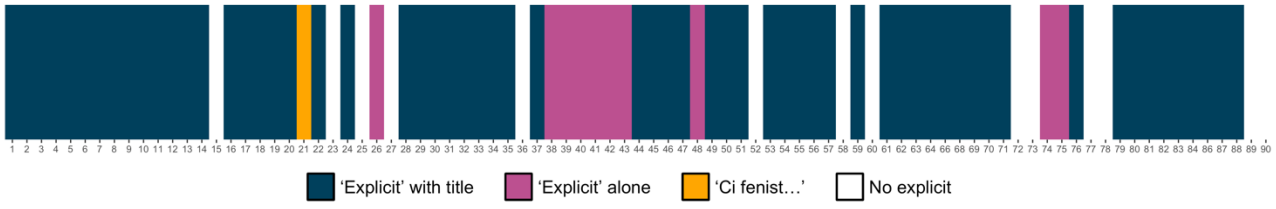
f. 441vb

Interestingly, the censure of sexual imagery carried out on f. 294ra (Figure 1.68) is also completed in red. This intervention could of course be attributed to the scribe responsible for the *litterae notabiliores* or any other individual with access to red ink, but if it were the rubricator who carried it out, it would indicate that this was not the same person as the scribe.

1C.3 *Explicit* markers

As we saw for *incipit* markers, the forms taken by the manuscript's *explicit* lines vary in a way that can help identify groupings of texts which may have come from the same exemplar. There are fewer *explicits* (73) than *incipits* (78), and unlike the markers used at the beginnings of text, these are always written in the same ink as the main text, usually by the same hand.

Figure 1.80 – visual representation of the distribution of different explicit *types within the manuscript.*[7]



What is notable from the visualisation in Figure 1.80 is that by far the default method of marking the end of a text is to include some variant of the word 'explicit', followed by the given title. 'Explicit' alone is not uncommon, however, and there is a single example in *Aristote* (item 21) of 'Ci fenist,' the Old French translation of the usual Latin. The *explicit* to *Respon*

[7] Again, I have excluded later additions from this visualisation.

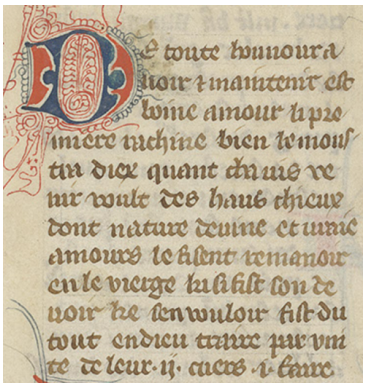
(item 3) is similarly interesting in that it attempts to make the verb reflexive as 'S'explicit,' thereby treating it as an Old French verb rather than a Latin one.

When compared with Figure 1.76, the cluster of texts in items 38–43 once again stands out as a group which may have been copied directly from a single exemplar manuscript. Interestingly, this exemplar may well have contained variation itself, since if we take items 36–45 to have been copied from the same manuscript, as is suggested by Figure 1.76, then the sub-group of items 38–43 may represent a cluster of texts that was originally together in one manuscript but which had been combined with other, comparable texts by the time they reached the exemplar from which they were copied into fr24432.

1C.4 Textual layout

The layout of the main body of the text in the manuscript is largely predictable, with the initials of verse lines usually separated from the rest of the line, and prose presented as a single block. Alexandrines are usually divided into hemistichs, while octosyllables and hexasyllables are presented with one orthographic line corresponding to one line of verse. Where metre is more variable, as in the various poems of *Serventois*, verse can be presented as prose with no separation made for the end of lines, except through apparent infrequent use of the *punctus*.

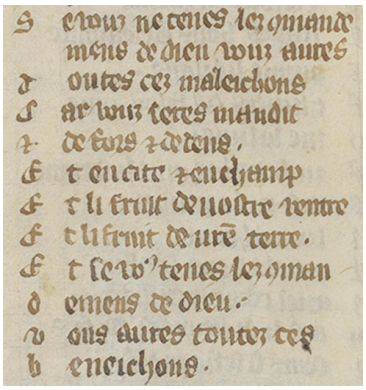
Figure 1.81 – the verse text Serventois *presented as prose, with a semi-diplomatic transcription (f. 306va).*



DE toute honnour avoir & maintenir
Est boine amour li premiere rachine
Bien le moustra Diex quant cha jus venir
Voult des haus chieux dont nature devine
Et vraie amours le fisent remanoir
En le Vierge ki si fist son devoir
Ke sen vouloir fist du tout en Dieu traire
Par unité de leur .ij. cuers .i. faire.

While this presentation of verse as prose is not uncommon,[8] what is more surprising is the apparent presentation of prose as verse. In quires 56 and 57, hand 11 appears to copy the verse layout in the preceding text (*Miroir*, item 88), shared between hands 2 and 11, even though the text being copied, *Salomon* (item 89), is in fact written in prose.

Figure 1.82 – the prose text Les Proverbes de Salomon *presented as verse, with a semi-diplomatic transcription (f. 430rb).*

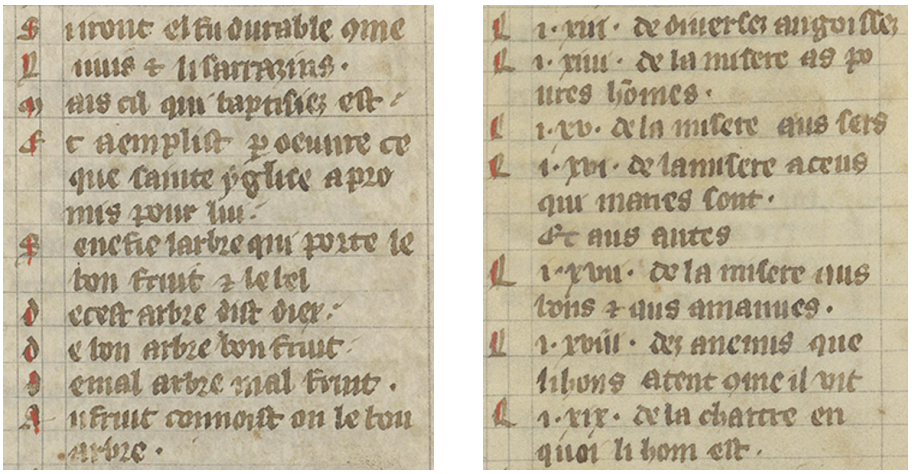


Se vouz ne tenés lez *com*mandemens de Dieu vouz avrés toutes cez maleichons car vouz serés maudit & defors & dedens. Et en cité & en champ et li fruit de vostre ventre et li fruit de v*ost*re terre. Et se vo*us* tenés lez *com*mandemens de Dieu; vous avrés toutez ces beneichons.

The layout becomes even more unusual on both sides of f. 436, found at the beginning of the following quire (q. 58), which contains the end of *Salomon*. Here, only initials at the beginning of a clause are permitted in the verse initial column. The best potential explanation for this layout is that the following folio (f. 437) contains a table of contents for the final text in the manuscript (*Misere*), which has a very similar layout to that found on f. 436, which may have influenced the scribe who copied this text (hand 11) in how they laid out the end of the text.

[8] This manner of presentation is particularly common in earlier manuscripts, especially those around the twelfth and thirteenth centuries. See, for example the copy of the *Vie de saint Alexis* found in Hildesheim, Dombibliothek, St Godehard MS 1, pp. 57–68. It is particularly common in Insular manuscripts, as for example the copy of Philippe de Thaon's *Bestiaire* in Oxford, Merton College Library, MS 249, ff. 1r–10v, perhaps on the influence of the concatenated layout standard for Old English verse. This could suggest that *Serventois* was copied from an exemplar which had retained an earlier approach.

Figure 1.83 – left: the unusual presentation of the end of Salomon *(f. 436ᵛᵇ); right: the table of contents for* Misere *from which the scribe appears to have drawn inspiration (f. 437ᵛᵇ).*



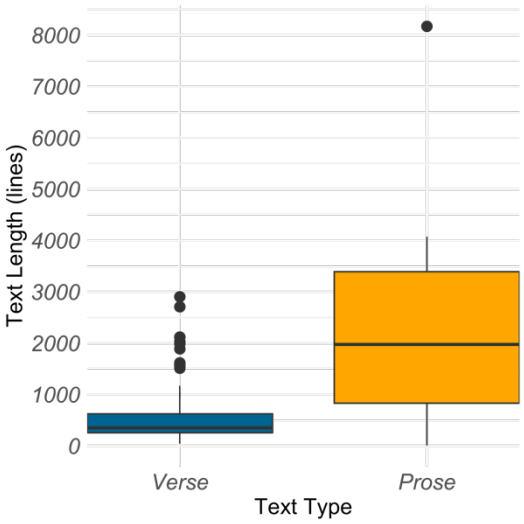
It is notable that these idiosyncrasies of layout do not appear to be simply the preference of individual scribes, since the continuation of *Misere* from f. 438 onwards, still in hand 11, returns to the default layout of presenting prose as a single block. While it is therefore possible that these apparent errors are due to decisions made by hand 11, it could also be the case that these errors were already present in the exemplar manuscripts from which the scribes copied the texts, which would reinforce the argument that *Salomon* and *Misere* came from the a single exemplar.

Appendix 1D: Further textual observations

1D.1 Form

Fr24432 contains 79 verse texts, eight prose texts, one mixed verse and prose text,[1] and a number of additions in later hands encompassing verse and prose.

Figure 1.84 – a box and whisker plot for the lengths of the manuscript's verse and prose texts in orthographic lines.



The prose texts are hugely diverse in length, with *Lucidaires*, the longest text of any kind in the codex at a little over 8,000 orthographic lines. This is more than double the length of the next longest prose text (*Aristote*), which places *Aristote* as a significant outlier within the data. This fact offers strong support for the argument, articulated throughout the main body of Chapter 1, that *Lucidaires* was not included in the compilation of the initial base manuscript, rather it was a later addition completed after the other texts had been copied but before the codex left the *scriptorium*.

Two of the prose texts also stand out by virtue of being unexpectedly short: *Dormant*, at just eight lines, and *Jorz*, at 42. However, both these shorter texts are lists rather than fully structured prose texts, and *Dormant* is even written in red ink rather than the brown ink used

[1] Rutebeuf's *Dit de l'erberie*. So as not to skew the data, I have excluded this, the fragmentary texts *Vilains*, *Vin*, *Lai d'amour 1*, and *Trois mortes* all later additions, and the two short lists *Dormant* and *Jorz* from Figure 1.84.

for other texts. This, combined with their vastly shorter length, suggests that both these texts were most likely considered distinct from other prose texts in the codex by the scribes and perhaps also early readers.

While their range is less than the prose texts, the manuscript's verse texts also vary quite significantly in length. The allegorical *Tournoiement* is the manuscript's longest verse text at nearly 3,000 lines, while *Despensier* is the shortest complete verse text at just 42 orthographic lines.

More than half of the verse texts are between 200 and 650 lines in length, and eleven appear to be outliers.[2] Unlike in the case of *Lucidaires*, however, it would not be accurate to claim that these longer texts are in some way aberrations from the norm, rather this variation in text length highlights that the manuscript compilator may well have had a general preference for shorter verse texts, while nevertheless also allowing longer ones of sufficient perceived value.[3]

1D.2 Versification and rhyme

The verse texts within the codex are diverse in form as well as length, with stanza structure, metre and rhyme all particularly variable. There are a total of 30 texts which consistently use alexandrines,[4] all of which are structured into monorhyme quatrains, with the exception of *Patenostres* (item 47), which appears in sexains that consist of a monorhyme quatrain followed by a two-line refrain with a consistent rhyme.[5]

[2] *Guillaume*, *Panthere*, *Isopet*, *Tournoiement*, *Robert*, *Florence*, *Anelés*, *Buef*, *Arbre*, *Serventois*, *Prise*. Outliers are defined here, assuming that the manuscript's verse and prose texts ought to form two normally distributed populations, as any value which satisfies either $L < Q_1 - 1.5(Q_3 - Q_1)$ or $L > Q_3 + 1.5(Q_3 - Q_1)$, where L is the text's length in orthographic lines.

[3] Cf. Chapter 3, in which we use text length as one of the metrics by which to compare corpora of texts which may or may not have shared the same author.

[4] Items 1, 3, 11, 20, 24, 26, 28–42, 44, 47, 53–57, 61, and 74. Given the volume of texts discussed, in this section I refer to texts predominantly by their item number, unless they are singled out.

[5] Naetebus (1891: 94) cites this as the only known example of the *aaaab*β rhyme scheme, with its static final line.

Octosyllabic verse texts, of which there are 38 in the codex,[6] are more variable in their stanza form and rhyme scheme. The vast majority (30) of these poems display irregular separation into stanzas, if such divisions are made at all, and are written in rhyming couplets, all of which is consistent with the use of this verse form for primarily narrative texts.[7]

Some exceptions do appear, however, with items 6, 19, 23, 59, and 81 preferring twelve-line stanzas, while *Charlot* (item 10) contains eight lines per stanza and *Trois mors* (item 2) has eighteen.[8] These texts are more variable in their rhyme schemes, with *Trois mors* appearing in rhyming couplets while items 6, 19, 23, 60, and 81 all display the mirrored rhyme scheme *aabaabbbabba*. *Divisions* (item 59) also uses this rhyme scheme, after a brief section of rhyming couplets in the prologue. *Charlot*, meanwhile, uses the alternating rhyme scheme *abababab*.

The remaining verse texts are more variable in their metre. A number of texts, generally categorised as a single item, reveal themselves to be composite texts in part through metrical shifts. *Nostre Dame* (item 48), for example, is a composite of three distinct texts on the same theme, with the first using alexandrine rhyming couplets and both others using octosyllables in the 'mirrored' rhyme scheme discussed above. Meanwhile, *Miroir* (item 88), a dream allegory, has a hexasyllabic prologue with the rhyme scheme *aabaab* which appears to have been added to the text later in the process of transmission, albeit likely before the present manuscript was compiled, which attaches to the beginning of the main text in octosyllabic couplets.

Item 50, the Paris II version of the *Ysopet*, contains 40 shorter texts, not including a prologue and an epilogue, which are not delineated in any way by the scribe who copied them, but do

[6] Items 2, 4–8, 10, 12, 16–19, 22–23, 43, 45–46, 49, 51–52, 58–60, 62–64, 66, 68–70, 75–76, 79–83, and 86–87.
[7] Items 4–5, 7–8, 12, 16–18, 22, 43, 45–46, 49, 51–52, 58, 62–64, 66, 68–70, 75–76, 79–80, 82–83, and 86–87.
[8] *Trois mortes* (item 60) appears likely to have been in twelve-line stanzas, given its use of the mirrored rhyme scheme, but this cannot be said with any certainty since only eleven lines of text survive.

appear to have been ordered according to their metre, with eight octosyllabic poems being placed first, followed by 32 hexasyllabic ones. All but the fourth, fifth, and sixth octosyllabic poems are written in non-strophic rhyming couplets, whereas the fourth text begins in quatrains with the rhyme scheme *abab*, before transitioning to sexains with the structure *aabaab*, the latter structure being that used in the following two poems. All of the hexasyllabic poems are organised into sizains which opt for an *aabccb* rhyme scheme, which in some stanzas ends up resembling an *aabaab* structure.

Some of the texts in fr24432 contain stanzas or repeating patterns in which the metre alternates, as is the case with items 5, 9, and 13 which repeat a rhyming triplet cluster of one line of four syllables followed by two octosyllables, although the very first line of the poem is usually an octosyllable.[9]

Directly following the second of these texts, at item 14, is *Traverces*, a short verse poem of a similar type which has a tail-rhyme structure, in which each line of octosyllable is followed by a line of four syllables. The rhyme scheme for this poem is intriguing in that the couplets appear to be offset and do not line up with the tail-rhyme structure, such that the first and last lines of the poem have no rhyming pair.

Item 73, an untitled *lai d'amour*, meanwhile, is particularly interesting since it makes use of thirteen-line stanzas which follow a highly unusual metrical structure: 8448844888448. It also has an unusual rhyme scheme which does not match up neatly with the metre: *aaabaaabbabba*.[10]

The remaining verse texts, *Serventois* and *Prise* (items 72 and 84), simply contain elements of metre which are unpredictably variable. *Prise* contains lyric insertions in the form of ballades

[9] The latter two of these texts are attributed to Rutebeuf.
[10] This combination of rhyme scheme and metre is the only such example observed by Naetebus (1891: 96).

and rondels with variable metre, although these tend to centre around the more predictable octosyllabic couplets that are used for the text's narrative elements. In all cases the lyric insertions begin with an alternating *aba* structure, but there are few other rhythmic similarities.

Serventois, meanwhile, perhaps placed close to item 73 on account of their non-standard structures, consists of a number of short texts of highly variable line length, centred around decasyllable.[11] Once again, the rhyme schemes of these poems usually begin with *aba*, with a great deal of subsequent variation.

Figure 1.85 – visual representation of the distribution of different metrical structures within the manuscript.

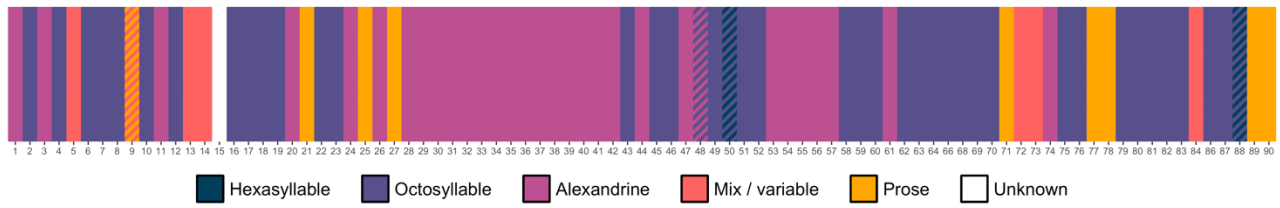


Figure 1.86 – visual representation of the distribution of different rhyme schemes within the manuscript.

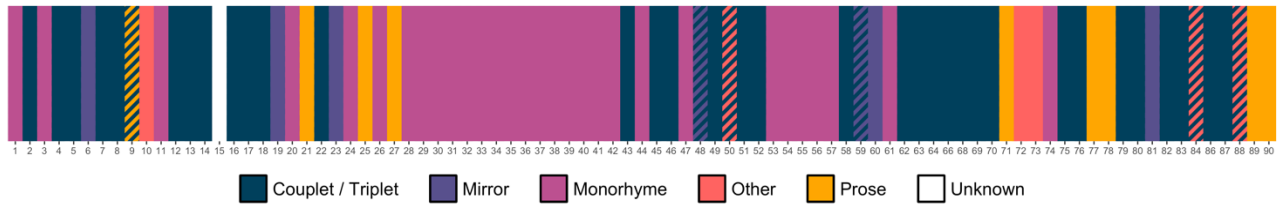


Figure 1.85 and Figure 1.86 both illustrate just how intertwined metre and rhyme scheme are within fr24432. They also reveal the fact that texts are interspersed throughout the manuscript, often appearing alongside texts of a different form, metrical structure, or rhyme

[11] It is possible that this could have influenced the scribe's decision not to separate the text into metrical lines, as displayed and discussed above for Figure 1.81, choosing instead to present the verse in the form of prose, perhaps as it was presented in the exemplar.

scheme.[12] That said, however, there are indications that poetic form may have been a contributing factor to the compilator's decision to place certain texts alongside each other.

In some cases, as in the long section of alexandrine quatrains at items 28–42, this juxtaposition of texts of the same form likely had much to do with the fact that their content is comparable, as well as perhaps to the possibility of shared authorship.[13] The similarities in metre and rhyme therefore possibly contributed to the sense of similarity that led the manuscript's compilator to place the texts together, but may not have been the primary motivating factor.

However, poetic form as an organising principle does appear to have played a crucial role in the juxtaposition of certain other texts. Items 13 and 14, as well as items 72 and 73, may have been placed near each other due to their complex and in some cases fully variable metre, since both pairs of texts share little else that would connect them directly.

The presence of the prose texts items 77 and 78, and 89 and 90 also indicates that prose form may have been considered sufficient reason to place one text next to another. Indeed, the fact that item 90 is incomplete raises the possibility that other prose texts may once have followed. As already argued in Chapter 1, *Lucidaires* (item 78) appears to have been a later addition to fr24432, and thus its inclusion after *Messe* (item 77) can be explained by the fact the recompilator wanted the text to fit into the broader structure of the manuscript, rather than feeling like an afterthought, hence the completion of the first text by hand 9, to avoid it becoming fragmentary.

[12] *Vilains* (item 15), left blank, survives only as an incipit, leaving no trace of the form it may have taken.
[13] As we argue in Chapter 3, all except item 29 (Rutebeuf's *Vie du monde*) within this group can be attributed to Jehan de Saint-Quentin, who asserts his status as author of *Chevalier et escuier* (item 33).

Appendix 3A: F-corpus contents

Item	Manuscript[1]	Short Title in Corpus	Full Title in Corpus	Attributed Author
1	Berlin 257 (C)	*Auberee*	*Auberee*	
2	Bern 354 (B)	*Berangier 1*	*Berengier au lonc cul (version 1)*	
3	BnF fr. 837 (A)	*Boivin*	*Boivin de Provins*	
4	Berlin 257 (C)	*BorgeoiseOrliens*	*La Borgeoise d'Orliens*	
5	BnF fr. 837 (A)	*BrunainBodel*	*Brunain la vache au prestre*	Jehan Bodel
6	BnF fr. 1593 (E)	*Crote*	*La Crote*	
7	Bern 354 (B)	*DaniuseleFoutre2*	*La Damoisele qui ne pooit oïr parler de foutre (version 2)*	
8	Bern 354 (B)	*DamoiseleSonjoit*	*La Damoiselle qui sonjoit*	
9	BnF fr. 837 (A)	*EstormiHuesPiaucele*	*Estormi*	Hues Piaucele
10	BnF fr. 1593 (E)	*FemmePendreMarieDeFrance*	*La Femme qui fist pendre son mari*	Marie de France
11	Bern 354 (B)	*Grue1*	*La Demoisele a la grue (version 1)*	
12	BnF fr. 12581 (X)	*MariEborgne*	*Le mari éborgné et sa femme infidèle (version 1)*	
13	BnF fr. 1630	*Nonete2*	*La Nonete (version 2)*	
14	Berlin 257 (C)	*quPrestresHaisiau*	*Les quatre Prestres*	Haisiau
15	BnF fr. 837 (A)	*Saineresse*	*La Saineresse*	
16	Bern 354 (B)	*SohaizDesvezBodel*	*Le Sohait desvez*	Jehan Bodel
17	BnF fr. 12603 (F)	*trAveuglesCortebarbe*	*Les trois Aveugles de Compeigne*	Cortebarbe
18	BnF Ars. 3525 (S)	*trDamesParisWatriquetDeCouvin*	*Les trois Dames de Paris*	Watriquet de Couvin
19	BnF fr. 1593 (E)	*trDamesVit1*	*Les trois Dames qui troverent un vit (version 1)*	
20	BnF fr. 25545 (I)	*unFemme100Chevaliers*	*Une seule Fame qui a son cors servoit cent chevaliers de tous poins*	
21	BnF fr. 19152 (D)	*VilainAsnier*	*Le Vilain asnier*	
22	Berlin 257 (C)	*VilainMire*	*Le Vilain Mire*	

[1] Manuscript sigla are given as in Noomen & van den Boogaard (1983–1998), with the exception of BnF fr. 1630, which was unknown to them. Full references for these manuscripts can be found in the bibliography.

Appendix 3B: Data

Table 3.47 – key statistics regarding the texts of the J-corpus.

Title	Length (Lines)	Prologue Length	Epilogue Length
Un dist que on clamme respon	192	12	7
Le dit des trois chanoinnes	412	16	4
Le dit des III pommes	336	8	4
Le dit de la bourjoise de Romme	216	8	8
Le dit des II chevaliers	210	13	6
Le dit de l'enfant rosti	180	7	4
Le dit du povre chevalier	160	8	4
Du chevalier et de l'escuier	228	4	11
Le dit de la borjoise de Narbonne	224	12	5
Le dit du chevalier qui devint hermite	172	16	4
Le dit du cordouanier	152	20	5
Le dit du petit Juitel	164	8	8
Le dit de l'enfant qui sauva sa mere	208	28	2
Le dit de l'eaue beneoite et du vergier	168	8	12
Le dit du riche home qui geta le pain a la teste du povre	168	12	8
Le dit du chien et du mescreant	204	20	16
Le dit de la pecherresse qui estrangla III enfans	152	12	8
Le dit de Merlin Mellot	224	14	12
De Florence de Romme	832	12	3
Le dit des anelés	780	84	4
Le dit du buef	796	40	8
Le dit de la beguine	204	32	7
La Vie saint Sauveur l'ermite	328	10	1
D'une abesse que Nostre Dame delivra de confusion	208	12	4

Table 3.48 – results of a p-value test for the use of le *in each text of the J-corpus, when compared to the rest of the corpus.*

le

Text	Obs Freq	Exp Freq[1]	p-value
Respon	26	23.2	0.52681
Chanoinnes	58	49.8	0.19806
Pommes	58	40.5	0.00260
Romme	39	26.1	0.00629
Deux chevaliers	27	25.4	0.72420
Enfant rosti	31	21.5	0.02655
Povre chevalier	13	19.3	0.12069
Chevalier et escuier	17	27.5	0.02813
Narbonne	35	27.1	0.09863
Chevalier hermite	21	20.5	0.91435
Cordouanier	27	18.4	0.03000
Juitel	17	19.8	0.49291
Enfant sauwa mere	26	25.1	0.85194
Eaue et vergier	17	20.3	0.43345
Pain	25	20.3	0.25924
Mescreant	24	24.6	0.89022
Pecherresse	10	18.4	0.03467
Merlin	25	27.1	0.67087
Florence	71	91.3	0.01660
Anelés	69	94.2	0.00339
Buef	79	95.9	0.04983
Beguine	25	24.6	0.94015
Abesse	36	25.0	0.01749
Sauveur	49	39.0	0.08013

Table 3.49 – results of a p-value test for the use of que *in each text of the J-corpus, when compared to the rest of the corpus.*

que

Text	Obs Freq	Exp Freq	p-value
Respon	13	14.2	0.73801
Chanoinnes	25	30.5	0.28752
Pommes	22	24.8	0.55709
Romme	11	16.0	0.18687
Deux chevaliers	19	15.5	0.35225
Enfant rosti	2	13.2	0.00164
Povre chevalier	14	11.8	0.50697
Chevalier et escuier	10	16.9	0.07590
Narbonne	10	16.6	0.08772
Chevalier hermite	11	12.6	0.64327
Cordouanier	4	11.2	0.02402
Juitel	11	12.1	0.73575
Enfant sauwa mere	12	15.4	0.36492
Eaue et vergier	17	12.4	0.17218
Pain	19	12.4	0.04900
Mescreant	21	15.1	0.10545
Pecherresse	13	11.2	0.58488
Merlin	16	16.6	0.88714
Florence	73	55.9	0.01255
Anelés	83	57.7	0.00037
Buef	58	58.7	0.91828
Beguine	12	15.1	0.40172
Abesse	17	15.3	0.64871
Sauveur	12	23.9	0.00985

Table 3.50 – results of a p-value test for the use of et *in each text of the J-corpus, when compared to the rest of the corpus.*

et

Text	Obs Freq	Exp Freq	p-value
Respon	10	13.6	0.30739
Chanoinnes	33	29.1	0.43899
Pommes	16	23.6	0.09186
Romme	13	15.2	0.54824
Deux chevaliers	12	14.8	0.43973
Enfant rosti	9	12.6	0.29080
Povre chevalier	4	11.3	0.02175
Chevalier et escuier	34	16.1	0.00000
Narbonne	12	15.8	0.31265
Chevalier hermite	13	12.0	0.76295
Cordouanier	15	10.7	0.17024
Juitel	13	11.6	0.66360
Enfant sauwa mere	20	14.7	0.14160
Eaue et vergier	14	11.9	0.51587
Pain	10	11.9	0.57421
Mescreant	31	14.4	0.00002
Pecherresse	6	10.7	0.12813
Merlin	13	15.8	0.45868
Florence	27	53.4	0.00013
Anelés	59	55.1	0.55901
Buef	68	56.0	0.07904
Beguine	23	14.4	0.01702
Abesse	13	14.6	0.65986
Sauveur	14	22.8	0.04984

[1] Observed Frequency: the number of observed instances. Expected Frequency: the number of instances that could be expected based on the rate across the corpus as a whole.

Table 3.51 – results of a p-value test for the use of de *in each text of the J-corpus, when compared to the rest of the corpus.*

de

Text	Obs Freq	Exp Freq	p-value
Respon	13	8.2	0.08094
Chanoinnes	19	17.6	0.72731
Pommes	8	14.3	0.07996
Romme	16	9.2	0.02020
Deux chevaliers	14	9.0	0.07944
Enfant rosti	6	7.6	0.54907
Povre chevalier	10	6.8	0.20927
Chevalier et escuier	3	9.7	0.02405
Narbonne	6	9.6	0.22722
Chevalier hermite	8	7.3	0.78306
Cordouanier	11	6.5	0.06360
Juitel	5	7.0	0.43444
Enfant sawa mere	8	8.9	0.75806
Eaue et vergier	7	7.2	0.94459
Pain	7	7.2	0.94512
Mescreant	6	8.7	0.33926
Pecherresse	10	6.5	0.15040
Merlin	10	9.6	0.88868
Florence	16	32.3	0.00181
Anelés	37	33.4	0.49276
Buef	43	34.0	0.09020
Beguine	8	8.7	0.80406
Abesse	6	8.9	0.32171
Sauveur	15	13.8	0.74309

Table 3.52 – results of a p-value test for the use of quand *in each text of the J-corpus, when compared to the rest of the corpus.*

quand

Text	Obs Freq	Exp Freq	p-value
Respon	6	8.0	0.47675
Chanoinnes	21	17.1	0.31856
Pommes	18	13.9	0.24559
Romme	4	9.0	0.08371
Deux chevaliers	5	8.7	0.19134
Enfant rosti	6	7.4	0.60377
Povre chevalier	5	6.6	0.51408
Chevalier et escuier	9	9.4	0.88429
Narbonne	11	9.3	0.56196
Chevalier hermite	9	7.0	0.45177
Cordouanier	4	6.3	0.34779
Juitel	15	6.8	0.00206
Enfant sawa mere	8	8.6	0.82977
Eaue et vergier	6	7.0	0.71318
Pain	8	7.0	0.68483
Mescreant	9	8.5	0.84766
Pecherresse	7	6.3	0.77455
Merlin	13	9.3	0.20711
Florence	25	31.3	0.22442
Anelés	24	32.3	0.11333
Buef	23	32.9	0.06100
Beguine	14	8.5	0.04521
Abesse	10	8.6	0.61995
Sauveur	23	13.4	0.00697

Table 3.53 – results of a p-value test for the use of a *in each text of the J-corpus, when compared to the rest of the corpus.*

a

Text	Obs Freq	Exp Freq	p-value
Respon	10	7.8	0.42866
Chanoinnes	30	16.8	0.00088
Pommes	15	13.7	0.71233
Romme	9	8.8	0.95202
Deux chevaliers	10	8.6	0.62100
Enfant rosti	8	7.3	0.78598
Povre chevalier	7	6.5	0.85948
Chevalier et escuier	7	9.3	0.43568
Narbonne	12	9.2	0.33122
Chevalier hermite	7	6.9	0.98298
Cordouanier	7	6.2	0.75146
Juitel	6	6.7	0.78241
Enfant sawa mere	8	8.5	0.86555
Eaue et vergier	3	6.9	0.12533
Pain	4	6.9	0.25801
Mescreant	5	8.3	0.23012
Pecherresse	4	6.2	0.36161
Merlin	9	9.2	0.96124
Florence	32	30.9	0.82727
Anelés	27	31.9	0.34959
Buef	32	32.4	0.93603
Beguine	8	8.3	0.90901
Abesse	4	8.5	0.11122
Sauveur	15	13.2	0.60497

Table 3.54 – results of p-value tests for the use of discourse onset Type A in each text of the J-corpus, when compared to the rest of the corpus.

Type A

Text	Obs Freq	Obs Rate	p-value
Respon	6	0.18	0.81030
Chanoinnes	11	0.34	0.03320
Pommes	11	0.24	0.48226
Romme	6	0.35	0.09988
Deux chevaliers	5	0.17	0.72299
Rosti	3	0.17	0.72691
Povre chevalier	4	0.17	0.69814
Chevalier et escuier	4	0.11	0.18226
Narbonne	5	0.21	0.91230
Chevalier hermite	6	0.24	0.60512
Cordouanier	3	0.15	0.59286
Juitel	9	0.45	0.00498
Enfant sauva mere	4	0.15	0.56491
Eaue et vergier	3	0.33	0.31026
Pain	4	0.25	0.64199
Mescreant	0	0.00	0.01501
Pecherresse	2	0.20	0.98954
Merlin	3	0.08	0.05961
Florence	10	0.15	0.29038
Anelés	11	0.19	0.80741
Buef	13	0.24	0.42474
Beguine	5	0.24	0.67099
Abesse	3	0.09	0.11476
Sauveur	15	0.28	0.12668

Table 3.55 – results of p-value tests for the use of discourse onset Type B in each text of the J-corpus, when compared to the rest of the corpus.

Type B

Text	Obs Freq	Obs Rate	p-value
Respon	11	0.33	0.71569
Chanoinnes	8	0.25	0.17415
Pommes	19	0.41	0.47735
Romme	10	0.59	0.05193
Deux chevaliers	16	0.55	0.02994
Rosti	11	0.61	0.02497
Povre chevalier	1	0.04	0.00071
Chevalier et escuier	19	0.53	0.03433
Narbonne	16	0.67	0.00148
Chevalier hermite	9	0.36	0.96870
Cordouanier	10	0.50	0.20547
Juitel	2	0.10	0.01222
Enfant sauva mere	9	0.35	0.85284
Eaue et vergier	4	0.44	0.62121
Pain	5	0.31	0.68313
Mescreant	11	0.48	0.24988
Pecherresse	2	0.20	0.27147
Merlin	3	0.08	0.00019
Florence	9	0.13	0.00006
Anelés	3	0.05	0.00000
Buef	23	0.43	0.32466
Beguine	11	0.52	0.12337
Abesse	23	0.70	0.00005
Sauveur	32	0.59	0.00021

Table 3.56 – results of p-value tests for the use of discourse onset Type C in each text of the J-corpus, when compared to the rest of the corpus.

Type C

Text	Obs Freq	Obs Rate	p-value
Respon	10	0.30	0.55422
Chanoinnes	12	0.38	0.78000
Pommes	9	0.20	0.02269
Romme	1	0.06	0.00915
Deux chevaliers	7	0.24	0.21007
Rosti	4	0.22	0.25587
Povre chevalier	7	0.29	0.53857
Chevalier et escuier	11	0.31	0.56090
Narbonne	3	0.13	0.01760
Chevalier hermite	9	0.36	0.92985
Cordouanier	5	0.25	0.34105
Juitel	5	0.25	0.34171
Enfant sauva mere	13	0.50	0.10678
Eaue et vergier	2	0.22	0.40717
Pain	7	0.44	0.48902
Mescreant	12	0.52	0.08150
Pecherresse	6	0.60	0.10533
Merlin	12	0.32	0.64024
Florence	48	0.72	0.00000
Anelés	44	0.75	0.00000
Buef	17	0.31	0.56001
Beguine	5	0.24	0.27617
Abesse	4	0.12	0.00446
Sauveur	5	0.09	0.00000

Table 3.57 – results of p-value tests for the use of discourse onset Type D in each text of the J corpus, when compared to the rest of the corpus.

Type D

Text	Obs Freq	Obs Rate	p-value
Respon	6	0.18	0.03452
Chanoinnes	1	0.03	0.27488
Pommes	7	0.15	0.08744
Romme	0	0.00	0.23053
Deux chevaliers	1	0.03	0.32975
Rosti	0	0.00	0.19407
Povre chevalier	12	0.50	0.00000
Chevalier et escuier	2	0.06	0.52034
Narbonne	0	0.00	0.12903
Chevalier hermite	1	0.04	0.41011
Cordouanier	2	0.10	0.86633
Juitel	4	0.20	0.05170
Enfant sauva mere	0	0.00	0.11752
Eaue et vergier	0	0.00	0.41059
Pain	0	0.00	0.26679
Mescreant	0	0.00	0.14450
Pecherresse	0	0.00	0.42204
Merlin	20	0.53	0.00000
Florence	0	0.00	0.01239
Anelés	1	0.02	0.05410
Buef	1	0.02	0.07151
Beguine	0	0.00	0.17516
Abesse	3	0.09	0.91963
Sauveur	2	0.04	0.19160

Appendix 3C: Scripts

Script 3.1 – R script to calculate a p-value for the use of a given feature between a target and reference corpus using randomised sampling.

```
#starting counts (tar = target corpus, ref = reference corpus)
observed_tar <- _ #observed instances of feature in target corpus
observed_ref <- _ #observed instances of feature in reference corpus

total_tar <- _ #total instances in target, with and without feature
total_ref <- _ #total instances in reference, with and without feature

sum_observed <- observed_tar + observed_ref #sum of observed instances
sum_total <- total_tar + total_ref #sum of total instances

#null rate: expected rate if there is no difference between a and b
null_rate <- sum_observed / sum_total

#observed difference between probabilities
observed_prob_dif <- (observed_ref / total_ref) - (observed_tar / total_tar)

#set sample size
total_n <- 1000000

#generate random samples using the null rate, counting observed instances
ref_rand <- rbinom(sample_size, total_ref, null_rate)
tar_rand <- rbinom(sample_size, total_tar, null_rate)

#work out difference in probability between two samples
rand_prob_difs <- (ref_rand / total_ref) - (tar_rand / total_ref)

#make differences in probability absolute
abs_rand_prob_difs <- abs(rand_prob_difs)
abs_observed_prob_dif <- abs(observed_prob_dif)

#calculate the p value (the proportion of the random samples for which the
    #difference in probability exceeds or equals the observed probability)
p_value <- mean(abs_rand_prob_difs >= abs_observed_prob_dif)

#return the p value
p_value
```

Appendix 3D: Codicological analysis

3D.1 Manuscripts

Aside from the digital and statistical analysis that we brought to bear in Chapter 3 in service of our examination of the shared author hypothesis, and coming back full circle to the kinds of analysis presented in Chapter 1, there is one final angle from which the poems of Jehan de Saint-Quentin may be considered, but thus far have not yet been. In particular, there is a good deal to be said about the ways in which Jehan's texts fit into the codicological structures of the manuscripts in which they appear.[1] Unlike the discussion in Chapter 3, which was directed towards the question of Jehan's authorship of the J-corpus, this appendix is broader in scope, attempting to provide an exhaustive analysis of the codicological contexts of Jehan's poems within their surviving manuscripts.

While the bulk of the texts of Jehan de Saint-Quentin's *œuvre* are to be found within fr24432, they are not exclusively limited to this one manuscript. As discussed in Chapter 3, there are two texts of the J-corpus which do not appear in the same manuscript as the others, namely *Sauveur*, which appears only in Paris, BnF Ars. 2115, ff. 48$^{r–v}$, 52$^{r–}$57^{v}, and *Abesse*, which appears only in Paris, BnF fr. 12483, ff. 66$^{ra–}$68^{rb}, the so-called *Rosarius* manuscript.

In addition to these copies of the texts, Morawski (1940: 519) and Munk Olsen (1978: cxlvii) record the presence of a copy of *Florence* within the fifteenth-century Grenoble, Bibliothèque municipale, 871, ff. 188$^{r–}$199^{r}, with Munk Olsen, the only of the two to include this text in his corpus, suggesting that this is the sole occasion on which a text of the J-corpus is found in more than a single codex. Both scholars appear to have missed a second copy of *Anelés*, which is to be found within the late-fifteenth-century codex Paris, BnF fr. 5036, ff. 216$^{r–}$225^{v}.[2]

[1] In this section I draw inspiration from the 'New Codicology' pioneered by Keith Busby (1999).
[2] Claude Roussel (2007), in the most recent work on *Anelés*, also suggests that the text is to be found only within fr24432. I was alerted to this second copy via the poem's entry in the *Archives de literature du Moyen Âge* (ARLIMA;

Achille Jubinal (1839–1842, ii: 421) and more recently Margherita Lecco (2020: 125, n. 3) and Gabriel Cholette (2022: 94) have claimed that *Merlin* can also be found within BnF fr. 12483, the so-called *Rosarius* dating from the early fourteenth century which contains *Abesse*. The tale included in the manuscript is in fact the version from the *Vie des peres* identified by Morawski (1939: 337) and Munk Olsen (1978: xcvii) as a likely source of *Merlin*, which is written in octosyllable rather than the alexandrine rhyming quatrains used in all of Jehan's poems.[3] It is still notable that *Abesse* should appear in the same codex as a putative source of Jehan's poems, suggesting that the texts were seen to fit into a pre-existing literary type.

Following our inclusion of *Guillaume* and *Robert* as texts which ought also to be attributed to Jehan, we ought also to note that in addition to its inclusion in fr24432,[4] *Guillaume* appears in an early fourteenth-century manuscript London, British Library, Additional 15606, making the copy in fr24432 likely not the earliest surviving copy of the text. *Robert*, meanwhile, also appears in a further three manuscripts, namely Besançon, Bibliothèque municipale, 588, Paris, BnF fr. 1881, and Paris, BnF fr. 12604. In this case, the manuscripts appear to have been produced a good deal later than fr24432.

Table 3.58 – the manuscripts containing copies of Jehan's texts, that is the J-corpus, Guillaume, *and* Robert.

Manuscript	Text Count	Century
Besançon, Bib. Mun. 588	1	Early 16th
Grenoble, Bib. mun. 871	1	15th
London, BL Add. MS 15606	1	Early 14th
Paris, BnF Ars. 2115	1	Late 15th
Paris, BnF fr. 1881	1	Late 15th / early 16th
Paris, BnF fr. 5036	1	2/2 15th
Paris, BnF fr. 12483	1	2/4 14th
Paris, BnF fr. 12604	1	15th
Paris, BnF fr. 24432	24	Mid-14th (after 1345)

Brun 2023; https://www.arlima.net/il/jean_de_saint-quentin.html#ann), but in the absence of any accompanying information, and no reference to the copy elsewhere, I have been unable to locate the discovery's original source.

[3] Jubinal does not give folio references, noting only that the text appears in chapter 26 of the *Rosarius*. Lecco interprets this as referring to the 26th quire (which consists solely of f. 258), and does not appear to have verified Jubinal's claim, stating that it is 'di difficile consultazione poiché composto da fascicoli non in ordine, in più misti di testo e musica.' The text in question is found on ff. 198ra–200vb. For an edition, see Méon (1823).

[4] An inclusion missed by Cholette (2022: 94).

In total, then, there are 9 manuscripts containing the works we have attributed to Jehan, presented in Table 3.58, of which only fr24432 contains more than a single attributed text. We note also that only BL Add. MS 15606, BnF fr. 12483, and fr24432 were produced before the fifteenth century, despite the fact that Jehan's texts were most likely composed in either the last quarter of the thirteenth century or the first quarter of the fourteenth, pointing to some degree of enduring popularity for Jehan's poetry in the centuries after its initial composition, even though it eventually fell out of the canon.

3D.2 Jehan's texts in fr24432

3D.2.1 The 'core' grouping

If we want to examine how Jehan's texts interact within a single manuscript, we are therefore only able to do so within fr24432. 14 of the 24 texts within fr24432 which we now attribute to Jehan form a single group, spanning ff. 99^{rb}–137^{rb}, with the only interruption offered by Rutebeuf's *Vie du monde*,[5] which appears after *Romme*, the first text of the group, on ff. 102^{rb}–104va. This run of texts, which falls squarely within the 'base' manuscript defined in Chapter 1, may be described as the 'core' grouping of Jehan's *œuvre*, since it contains more than half of the 26 texts which we have attributed to Jehan found in any manuscripts.[6]

Table 3.59 – texts attributed to Jehan in the core grouping in fr24432.

Text	Item	from fol.	to fol.
Romme	28	99^{rb}	102^{rb}
Deux chevaliers	30	104va	107^{rb}
Enfant rosti	31	107^{rb}	109^{vb}
Povre chevalier	32	109^{vb}	112ra
Chevalier et escuier	33	112ra	115^{rb}
Narbonne	34	115^{rb}	118^{rb}
Chevalier hermite	35	118^{rb}	120va

[5] For an edition of this version, see Bastin & Faral (1959–1960, i: 395–407).

[6] That is, the 24 texts of the J-corpus, in addition to *Guillaume* and *Robert*. I consciously use the term 'grouping' here as opposed to Busby's preferred term 'sequence,' since this appears to be a better fit for fr24432. The compilation does not feel especially linear, with texts placed in a particular order to be read in sequence. Rather, it is as if texts were seen by the compiler to hold a sort of thematic, linguistic, and structural gravity, attracting other, similar texts into a particular orbit, thus forming a grouping of texts which has both internal connections, and connections with nearby texts external to the grouping, but which is not intended to be read in a linear way from start to finish.

Cordouanier	36	120va	122vb
Juitel	37	122vb	125ra
Enfant sauva mere	38	125ra	128ra
Eaue et vergier	39	128ra	130rb
Pain	40	130rb	132va
Mescreant	41	132va	135rb
Pecherresse	42	135rb	137rb

This grouping spans several quires (qq. 13–18), but notably neither starts from the beginning of quire 13, nor finishes at the end of quire 18. The first text by Jehan in the grouping is written by one scribe (hand 5), who also starts the *Vie du monde* before handing over to a second scribe (hand 6) who completes the grouping, continuing on until the end of quire 26. As such, the grouping appears to have been conceived as part of the broader continuation of the manuscript's compilatory structure, rather than having been in any way *mis à part*.

We may well attribute the grouping of these texts to the thematic similarities which they contain. All are concerned primarily, or indeed exclusively, with questions of religious doctrine and behaviour, and of the performance of miracles. Mary plays a particularly prominent role, appearing directly in no fewer than 10 of the 14 texts,[7] while the Devil himself appears in the same number.[8] The texts are therefore largely concerned with questions of sin and salvation, which could explain the apparent intrusion of the *Vie du monde* into the group, as an estates satire lamenting the sins of the world and the misery that results.

These thematic similarities therefore appear to have motivated the juxtaposition of these texts within the group, but perhaps the most important motivating factor behind the decision to place the group of texts in this section of the manuscript appears to be poetic form. As we saw in Chapter 3, all of Jehan's texts are written in four-line stanzas of alexandrine monorhyme, and the *Vie du monde* also fits this description. The text is thus an outlier in the sense that it is

[7] *Romme*, *Deux chevaliers*, *Povre chevalier*, *Chevalier et escuier*, *Narbonne*, *Chevalier hermite*, *Cordouanier*, *Enfant sauva mere*, *Eaue et vergier*, and *Pecherresse*.
[8] *Romme*, *Deux chevaliers*, *Enfant rosti*, *Povre chevalier*, *Chevalier et escuier*, *Narbonne*, *Chevalier hermite*, *Cordouanier*, *Enfant sauva mere*, and *Pecherresse*, thus giving an overlap of 9 texts which feature both Mary and the Devil.

not a text written by Jehan, but by many other metrics it is as valid a candidate for inclusion in this group as any of Jehan's texts.

Ignoring two very short intervening prose lists which appear to have been opportunistically placed in an appropriate gap between two texts,[9] we find that this 'core' grouping extends earlier than the first of Jehan's poems. There are in fact two preceding texts which also fit the criteria for inclusion, namely *Bible* (ff. 87^{rb}–91^{rb}) and *Vision* (ff. 91va–99ra).[10] The former of these comprises praise of Mary through moralising explanations of elements of the Old Testament which supposedly prophesise her life, while the latter is a vernacular version of the *Visio sancti Pauli*. Both therefore have explicitly didactic, religious content, and as we saw from their inclusion in the P1a-corpus alongside the *Vie du monde*, both are written in alexandrine monorhyme quatrains, reinforcing the likelihood that these features were the motivating factors behind the compiler's organisation of this particular section of fr24432.

Accepting these two earlier texts into our grouping, we may therefore say that the 'core' grouping containing Jehan's texts extends all the way back to the first folio of quire 12, itself the work of a third scribe (hand 3). The contiguous nature of the interaction between this grouping and the rest of the manuscript is once again made clear, however, by the fact that the beginning of the grouping does not exactly correlate with start of a new quire, since earlier on this first folio of quire 12 we find the end of *Espervier* (ff. 86va–87ra), itself an octosyllabic invective against avaricious, inconstant women, which is also copied by hand 3. This text perhaps acts as a counterpoint to the praise of Mary that follows, but it certainly does not appear to be part of the grouping containing Jehan's texts, rather more likely forming part of a previous grouping of texts concerned with behaviour within society, as opposed to religious behaviour, which spans ff. 57ra–87ra, and is dominated by *Aristote*, a

[9] *Dormant* (f. 91va), and *Jorz* (ff. 99ra–99^{rb}).
[10] For an edition of the former text, see Wilmotte (1886), and for the latter, see Kastner (1906).

vernacular translation of the pseudo-Aristotelian *Secretum secretorum*, which lasts from the beginning of the grouping until f. 84ra.

At the other end of the 'core' grouping, we find a few texts soon after *Pecherresse* which are written in alexandrines,[11] which we may choose to include in our grouping, but the thematic similarities are less clear for the earlier of these texts, suggesting some kind of rupture. This sense of separation is also reinforced by the presence of a few intervening poems which differ more or less significantly from the other texts in the grouping in their content, and all of which are distinct in their form, being written in the octosyllabic rhyming couplets common elsewhere in the manuscript.[12] Where the beginning of the grouping is clear, then, the texts from ff. 137rb–153rb, by contrast, may constitute some sort of bridging section, since the compilation then turns to questions of love and romantic relationships, in particular through the inclusion of the long dream allegory *Panthere* (ff. 153va–171rb).

While thematic and formal factors therefore appear to have played a role in the 'core' grouping, authorship appears to have been less important, since the manuscript's rubrication makes no attempt to identify the texts' authors within the main body of the grouping, despite the appearance within the grouping of *Chevalier et escuier*, the text in which Jehan gives his name. In the bridging section at the end of the grouping, the author of *Bacchus* is given as 'Gieffroy' in both the *incipit* and *explicit*, while a date of composition is given in the *explicits* each of *Mais* (1324), *Bacchus* (1313), and *Patenostres* (1320), but this does not appear to give any clues as to the reasons for the texts' compilation. If anything, since the scribe for this bridging section is the same as earlier in the grouping (hand 6), this would appear to suggest that the

[11] *Mais* (ff. 142rb–146va), *Patenostres* (ff. 148rb–152ra), and the first of the three sections of *Nostre Dame* (ff. 152ra–153rb). *Mais*, as we have seen, is written in monorhyme quatrains like the rest of the grouping, while *Patenostres* appears in quatrains with an accompanying two-line refrain, and the initial section of *Nostre Dame* is written in rhyming couplets in twelve-line stanzas, before the text switches to octosyllables.

[12] *Rebellion* (ff. 137rb–138va, anti-English and anti-Flemish invective presented as advice to a king), *Bacchus* (ff. 142vb–146va, a parodic hagiography of Saint Bacchus, that is, of wine), and *Requeste* (ff. 146va–148rb, a satirical criticism of the monastic orders).

difference in rubrication likely stems from the fact that the texts of the bridging section were copied from a separate exemplar manuscript to that used for the main body of the section, supporting an interpretation that separates these later texts from those attributed to Jehan, but not for direct reasons of authorship: the exemplar from which Jehan's texts were copied may have borne closer resemblance to Jehan's autograph manuscript, if one ever existed, but this does not mean that the compiler of fr24432 believed, or indeed cared, that the texts he grouped together in this section were the work of a single poet.

3D.2.2 The 'long-text' grouping

As discussed in our examination of text length from page 151, it is also possible to identify an additional group of Jehan's texts within fr24432 consisting of *Guillaume*, *Robert*, *Florence*, *Anelés*, and *Buef*. All of these texts except for *Guillaume* appear in a similar location within fr24432, forming a grouping which we may term the 'long-text' grouping. This section of the codex comes later in the manuscript than the 'core' grouping, spanning from f. 199ra to at least f. 241^{rb}, although it could arguably be extended further. It begins with *Merlin* (ff. 199ra–201^{vb}), which is not itself particularly long in comparison to the rest of Jehan's texts, at only 224 lines, but it marks the beginning of a section containing a number of longer texts written in alexandrine monorhyme quatrains, including those which we have attributed to Jehan. *Merlin* is followed immediately by *Robert* (ff. 202ra–215^{rb}), which we have seen is deserving of inclusion amongst Jehan's *œuvre* on a wide array of evidence, even though it has thus far largely been ignored by scholarship. *Robert* is the longest text of those we have attributed to Jehan, at some 1,016 lines, and it is followed immediately by *Florence* (ff. 215^{rb}–226va), which at 832 lines is the longest text of the J-corpus of texts attributed to Jehan by previous scholarship (although 76 of these are missing in fr24432, due to the loss of f. 223).

Florence is followed by a shorter poem not attributed to Jehan, *Menage* (ff. 226va–230^{vb}), which we have seen to differ in its content from the surrounding texts, being a primarily dialogic text

in which three speakers debate the merits, or otherwise, of marriage and priesthood. However, the connection of this poem to the 'long-text' grouping comes again from the fact that it is written in alexandrine monorhyme quatrains, which reinforces the apparent importance of form in the organisation of this section of the manuscript, and the fact that, at 328 lines, it is amongst the longer alexandrine texts observed in fr24432. Following *Menage*, we find *Anelés* (ff. 231ra–241^{rb}), another text by Jehan, which at 780 lines, is also amongst his longest, and which therefore also makes sense as a text to be included in this grouping.

Table 3.60 – texts found in and after the 'long-text' grouping in fr24432. Those texts in bold are those which we have attributed to Jehan.

Title	from fol.	to fol.
Merlin	199ra	201^{vb}
Robert	202ra	215^{rb}
Florence	215^{rb}	226va
Menage	226va	230^{vb}
Anelés	231ra	241^{rb}
Yonec	241^{rb}	245ra
Divisions	245ra	246^{vb}
Trois mortes	246^{rb}	—
Buef	247ra	257va

Again, this section of the 'long-text' grouping spans both quires and scribes, with the texts from *Merlin* to *Menage* being the work of one scribe (hand 7) across quires 27 to 30, while *Anelés*, which fills quire 31 and spills over onto the first three leaves of quire 32, is the shared work of two scribes (hand 8 for quire 31, followed by hand 3 from the beginning of quire 32).

What differentiates the 'long-text' grouping from the 'core' grouping, however, is that it appears to have a discrete beginning, and is not contiguous with the preceding contents of the manuscript. The previous text in the codex, an untitled *lai d'amour* in octosyllabic rhyming couplets, is fragmentary, appearing only in the two columns of f. 198G^{v}. The quire in which the text appears (q. 26) is complete, suggesting that at least one folio has been lost here from the original compilation. Given the other evidence for continuity between qq. 26 and 27 discussed in Chapter 1, predominantly the continuous contemporary foliation, it appears

likely that *Merlin* was indeed the first text in this grouping, and that that the 'long-text' grouping was conceived as a distinct section from the texts which had originally come before it.

As we argued in Chapter 1, the section from ff. 231–246 (qq. 31–32) appears to be a later inclusion in fr24432, albeit one originally produced in the same scriptorium, which was inserted within the originally contiguous 'base' manuscript, in between *Menage* and *Buef* which originally appeared side by side. This is significant for the purposes of understanding the 'long-text' grouping, since it allows us to explain the inclusion of texts which appear out of place in the grouping. *Yonec*, *Divisions*, and *Trois mortes* have no clear reason for inclusion in the grouping, being short octosyllabic poems which do not fit the themes of Jehan's texts, and yet the fragmentary nature of *Trois mortes* in particular, reinforced by its exclusion from the table of contents,[13] indicates that these shorter texts ended up in the 'long-text' grouping by chance, rather than by design. *Anelés* spans the two inserted quires, and therefore the recompilation, and the inclusion of quires 31 and 32 at this point in the manuscript, can be explained by a desire on the part of the compiler to include *Anelés* alongside the other texts of the 'long-text' grouping, placing it in the obvious location for such an insertion opened up by the fact that *Menage* finishes at the end of one quire, and *Buef* begins the next. The shorter texts were therefore included for material rather than textual reasons, and thus the integrity of the 'long-text' grouping appears unaffected.

On this understanding of the recompilation processes affecting fr24432, the significance of this grouping becomes even more apparent. The organising principle for this section was the inclusion of long texts with didactic content, written in alexandrine monorhyme quatrains, to such an extent that the compiler moved material in from elsewhere to complement what had

[13] On which, see the discussion on p. 47.

already been copied. As with the 'core' grouping, it therefore appears undeniable that there has been an attempt to group texts by Jehan together, but this is not on the basis of his asserted authorship of the texts, given the inclusion of *Menage*, rather it is their formal and thematic similarity which makes them of interest to the compiler.

3D.2.3 Single texts elsewhere in fr24432

Outside of the two main groupings, there are five texts attributed to Jehan within fr24432 which do not appear in clear groupings of texts including others of his poems. The first of these is *Guillaume*, which as we have seen is a text which has not traditionally been attributed to Jehan, but which has all the necessary similarities with the texts of the J-corpus for such attribution to be possible. The poem tells the story of the divine trials of Guillaume, king of England, and appears on ff. 1^{ra}–13^{va}. It is thus the first text of any kind to appear in fr24432, and is copied by a single scribe (hand 2), spanning roughly one and a half quires (qq. 1–2).

Only a few folios after *Guillaume* we find *Respon* (ff. 14^{rb}–17^{ra}), a story about a young girl who is saved from rape and servitude by the fact that her name is Mary. It also appears entirely within the work of the same scribe as *Guillaume* (hand 2), and is mainly within a single quire (q. 2). Its final column appears at the beginning of quire 3, but there is no evidence to suggest that this break occurs for any other reason than scribal necessity. *Guillaume* and *Respon* are thus comparable in their form, and in the fact that they tell stories of individuals who have been saved by piety, but it is notable that none of these features are shared by the texts which immediately surround the texts, making the presence of a coherent group at this point in fr24432 appear unlikely.

Both *Trois mors* (ff. 13^{va}–14^{rb}), which follows *Guillaume* and precedes *Respon*, and *Mesdisans 1* (ff. $17^{ra}22^{vb}$), which follows *Respon*, are written in in octosyllabic rhyming couplets, rather than in alexandrines. Neither share a great deal of thematic similarity with either *Guillaume* or *Respon*,

since the former is a short dialogic *memento mori* depicting male characters (and therefore not to be confused with the fragmentary *Trois mortes* (f. 246^{rb}), which depicts female ones), while the latter is a love allegory which appears twice in the manuscript, both here and later on ff. 407^{rb}–412^{rb}.

There is therefore no clear reason why *Guillaume* and *Respon* would appear in close proximity in fr24432, especially alongside *Trois mors* and *Mesdisans 1*. The decision to begin fr24432 with *Guillaume* may have been motivated by a specific structural reason, or simply on the basis of which text was available first. In turn, *Respon* may have been included two texts later either by choice, as a result of its thematic and formal similarity, or by the simple fact that the text appeared in a comparable place in the exemplar manuscript, in which case the compilatory decision can be passed further back up the chain of transmission.

Turning to *Chanoinnes* (ff. 36^{vb}–42ra), the next text by Jehan to appear outside of a clearly defined grouping, we find a story about how three wealthy canons, having converted to hermitic life, help a foolish knight be redeemed of his sins. It is the work of the same scribe behind *Guillaume* and *Respon* (hand 2), and spans quires 5 and 6, again forming part of a contiguous progression of texts within the 'base' manuscript unit of fr24432. The poem is preceded by two tongue-in-cheek works by Rutebeuf, *Erberie* (ff. 33^{vb}–35va), and *Charlot* (ff. 35^{vb}–36^{vb}), of which the former is mixed-form, beginning with a 488 metrical structure and ending in prose, and the latter is more conventionally composed in octosyllables with an alternating rhyming pattern. Afterwards comes *Oiselet* (ff. 42ra–45^{rb}), a poem on the wickedness of *vilains*, itself written in octosyllabic rhyming couplets, which is then followed by another text by Rutebeuf, *Complainte* (ff. 45va–46va), which has the same 488 metrical structure as the first section of *Erberie*.

On the face of things, then, there are no formal or thematic similarities to indicate why a compiler would choose to include *Chanoinnes* in amongst these texts. One possible explanation, however, could come from the fact that a text apparently by Rutebeuf, the *Vie du monde*, was considered worth including alongside Jehan's texts in the 'core' grouping, which may also have influenced the compilation of *Chanoinnes* alongside other texts by Rutebeuf, suggesting that the compiler saw some element of similarity between Rutebeuf's texts and those of Jehan. If this was indeed the motivating factor for the order of texts at this moment in the codex, then this could also be the first example of a compilatory process which pays at least some attention to authorial attribution. This hypothesis can only be made tentatively, but the possible importance of authorship here is supported by the fact that Rutebeuf's authorship of a number of these texts is asserted in the manuscript's paratexts, with *Erberie* and *Complainte* attributed to Rutebeuf in both their *incipits* and *explicits*.[14]

Later in fr24432, from ff. 53va–56^{vb}, we find *Pommes*, another of Jehan's texts which, in ways reminiscent of the story of *Ami et Amile*, tells of two men who develop a deep friendship while on pilgrimage, and who each rely on the other in their hour of greatest need. None of the texts immediately surrounding *Pommes* bear any resemblance in form or content, since the preceding text is an elaboration on the *Ave Maria* in octosyllabic stanzas of twelve lines with the unusual 'mirror' rhyme scheme *aabaabbbabba*, and the text which follows is *Aristote*, the prose vernacular translation of the pseudo-Aristotelian *Secretum secretorum*, which begins the section of texts on behaviour within polite society that immediately precedes the 'core' grouping of Jehan's poems, as discussed above.

[14] For *Erberie*, the *incipit* reads 'Ci commance l'erberie Rustebuef' (f. 35^{vb}) while the *explicit* reads 'Explicit l'erberie Rustebuef' (f. 35va). In *Complainte*, the paratexts refer to the texts by its alternative title *L'œil Rutebeuf*, with the *incipit* reading 'Ci commance le dit de l'ueil Rustebuef' (f. 45va), and the *explicit* 'Explicit le dit de l'ueil Rustebuef' (f. 46va).

Why, then, could *Pommes* have been placed in this position? A possible explanation comes from the quire and scribal structure, since *Pommes* appears at the very end of quire 7, and is the last text copied by hand 2 in this particular section of the manuscript, with another scribe (hand 3) taking over from the beginning of quire 8. While both quires fit within the contiguous 'base' manuscript of fr24432, *Pommes* may well have been included simply because it fitted the available space before the end of quire 7, after which point the compiler wanted to begin a new quire with a new section of texts. *Pommes* may therefore have been drawn from an exemplar manuscript which also have contained *Guillaume*, *Respon* and *Chanoinnes*, but was placed away from them solely for reasons of material convenience.

The final of Jehan's poems to appear in fr24432 without being associated with any grouping of his texts comes much later, on ff. 312^{vb}–315va. This text, *Beguine*, is a Eucharistic miracle about a beguine who attempts to violate the sanctity of the Host for personal gain, but who is prevented from succeeding when the body of Christ cries out like a new-born child.

Again, there are no other alexandrine texts in sight here, and in this case the text all fits alongside other texts within a single quire, which is itself the work of a single scribe (hand 5), and again sits within the 'base' manuscript of fr24432. Before the poem is another untitled *lai d'amour* (ff. 312^{vb}–315va), which bears little thematic similarity to *Beguine*, but the texts which follow it do appear to explain its presence at this point in the manuscript. The text is followed immediately by *Anel* (ff. 315va–316^{vb}), one of Gautier de Coinci's *Miracles de Notre Dame*, in which a novice, having promised to dedicate his life to Mary, fails to do so and is reminded of his promise by the presence of the Blessed Virgin on his marital bed. The text is formally distinct from *Beguine*, being written in rhyming octosyllabic couplets, but shares a great deal of thematic similarity with it, in that both discuss the dangers and consequences of breaking a sacred oath. Both texts thus appear to have been included as a precursor to the much longer *Messe*, a vernacular translation of Hugues de Saint-Cher's *Speculum ecclesiae* (ff. 317^{rb}–327^{rb}),

which contains explanations of the significance of the rites of mass, including the giving and receiving of the Host, and appears after a brief humorous *excursus* in the form of *Cheval* (316^{vb}–317^{rb}), in which a horse trader boasts about the terrible quality of the animal he has for sale.

Within fr24432, we can therefore identify three main motivating factors behind the places in which Jehan's texts are compiled: form, content, and authorial association. The first two of these factors are by far the most cogent, such that we cannot use even the major groupings of Jehan's texts as concrete evidence for shared authorship. Nevertheless, examination of Jehan's texts in fr24432 is enormously revealing both of the practices of compilation and recompilation that took place in the manuscript's production, as well as what was considered important about the texts in the decades following their composition.

3D.3 Jehan's texts in other manuscripts

Having examined the motivating factors behind the inclusion and positioning of Jehan's texts in fr24432, let us now turn to the eight other manuscripts in which his texts are found.[15] We begin with those texts which are found in fr24432 and repeated elsewhere, and then continue with those which have been attributed to Jehan but do not appear in the same codex as any other of his texts. Note that since the focus here is codicological, rather than editorial, we will not consider textual variation when more than one copy of a given text exists, instead we will examine the codicological contexts of each of Jehan's poems when they appear outside of fr24432.

3D.3.1 Guillaume *in BL Add. MS 15606*

BL Add. MS 15606 is apparently the oldest manuscript to contain any of Jehan's work. Dating to the early fourteenth century, it was therefore compiled reasonably soon after many of Jehan's texts were written, assuming that this took place between the latter decades of the

[15] In all cases, I have begun my investigations on these manuscripts from the ARLIMA database.

thirteenth century, and the first decades of the fourteenth. It is therefore enlightening to see the kinds of texts which *Guillaume* appears alongside in this early version. The manuscript consists of 162 leaves, and contains some 32 items, written primarily in Old French. These include *Des Poignes d'enfer*, the apocryphal tale of Saint Paul's visit to the underworld which also appears in fr24432 (*Vision*, ff. 81ra–87^{vb}). Like fr24432, BL Add. MS 15606 also contains a prose treatise on the mass, and there is a definite interest for the compiler(s) in texts which contain doctrinal or moral teachings, made evident through the presence of texts such as *De David le prophecie* (ff. 6ra–17^{vb}),[16] an allegorical battle between vices and virtues written in octosyllabic rhyming couplets which is reminiscent of Huon de Méry's *Tournoiement* on ff. 184ra–198G^{rb} of fr24432.

Unlike fr24432, however, the manuscript also contains elements taken directly from scripture and the liturgy, including a copy of the *Gloria in excelsis*, the *Credo*, and the beginning of Saint John's Gospel on f. 157, as well as a French translation of the penitential psalms on ff. 97va–100ra. Similarly, there is also a sermon (ff. 90va–96^{vb}), and several prayers to both Jesus and Mary which largely appear in sequence (ff. 87^{vb}–90ra).

Guillaume fits comfortably into BL Add. MS 15606, appearing on ff. 140va–151^{vb}, immediately after *Des apostres qui anoncient la loi nostre signour Jhesu Crist* (ff. 130ra–140^{rb}), a prose version of the life of Saint Denis of Paris.[17] It is therefore likely that *Guillaume* was included in the manuscript on account of its somewhat hagiographic description of the life of a pious individual, which was deemed to fit alongside the preceding text. Formal features were apparently of lesser importance, given that *Des apostres* is written in prose, and *Guillaume* is followed immediately afterwards by another prose text, specifically *Li livres de sapience* (ff.

[16] For an edition, see Fuhrken (1895).
[17] This text is begun twice in the manuscript, appearing also on ff. 161–162, but this copy is fragmentary. For an edition of the text, as well as others in the same manuscript, see Meyer (1877).

152ra–156va), which presents the wisdom of Biblical prophets and is therefore comparable to *Salomon*, another prose text which appears in fr24432 on ff. 420ra–436vb.

BL Add. MS 15606 therefore shares a number of points of comparison with fr24432, but differs in its inclusion of scriptural texts. Similarly, the British Library manuscript does not contain as many short narrative texts as fr24432, with the compiler(s) preferring to include texts with a more direct moral or doctrinal message. While the inclusion of *Guillaume* in this manuscript is therefore explicable on the basis of its didactic content, the poem is a good deal more peripheral than it is in fr24432, a fact which foregrounds the significance of narrative texts within fr24432. Through its inclusion alongside less narrative didactic texts, *Guillaume* therefore takes on a function more akin to that which we have discussed for *exempla*, as a text inserted among others with more explicit moral messages, in order to illustrate the points contained within them.

3D.3.2 Robert *in BnF fr. 12604, BnF fr. 1881, and Besançon Bib. mun. 588*

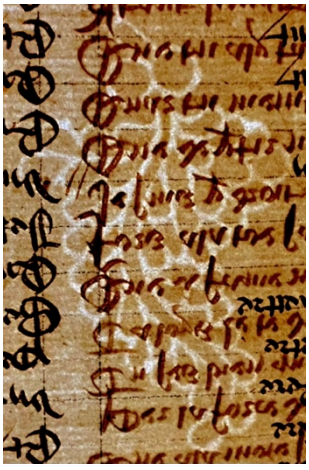
BnF fr. 12604

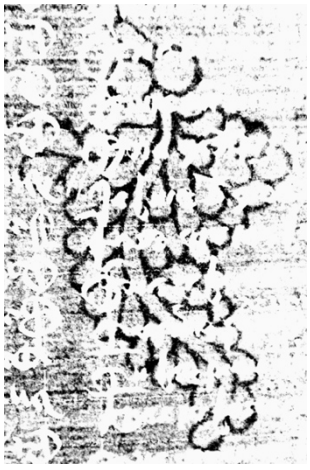
Robert, being the only text which we have attributed to Jehan to appear in more than two manuscripts, allows us to take an even more comparative approach. Beginning with BnF fr. 12604, a fifteenth-century paper manuscript containing 125 leaves, we find that the codex contains only two distinct texts, namely the *Vie des peres* (ff. 1ra–110rb), followed by *Robert* on ff. 110va–125ra). While the manuscript is therefore not a particularly complex one in terms of its compilation, the inclusion of *Robert* alongside the *Vie des peres* is especially significant, given that as we have seen the latter text is one which contains tales closely related to a number of Jehan's poems, indeed being referenced by him in *Mescreant* as a source of his text. It is therefore all the more significant that *Robert* is introduced within the manuscript as 'La vie de Robert le dyable,' suggesting that the text was taken to be a logical continuation of the tales in

the *Vie des peres* on account of its content, despite the fact that both texts are written in differing metres, with the *Vie des peres* appearing in octosyllables, and *Robert* in alexandrines.

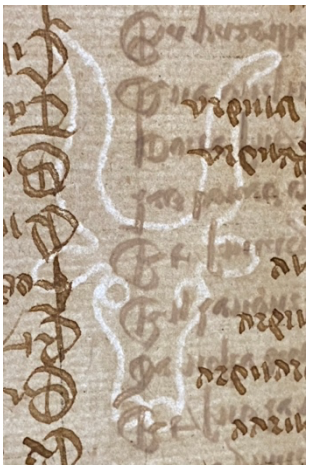
Robert is copied in a different hand to the preceding text, and in a lighter ink, as is the entry of the text into the table of contents. It would therefore be tempting to suggest that *Robert* was included some time after the *Vie des peres* was copied, perhaps even after the codex had moved to a new owner, who chose to include *Robert* on the basis of its textual and thematic similarities with the *Vie des peres*. However, there is evidence to refute this, which stems from the watermarks found in the paper on which both the *Vie des peres* and *Robert* were copied. The beginning of the codex displays consistent use of a watermark resembling a bunch of grapes, as displayed in Figure 3.87. From f. 23, this switches to a bull's head, before the grapes reappear from f. 120 to the end.[18] The shifts in watermarks therefore do not align with a shift in text, and *Robert* and the *Vie des peres* are both written on paper displaying both watermarks. It is therefore likely that *Robert* was indeed added to the codex later, on account of its textual and thematic similarities to the *Vie des peres*, but that it was copied on the same stock of paper, likely therefore by the same individual.

Figure 3.87 – watermarks found in BnF fr. 12604 (top: f. 22; bottom: f. 24).





[18] The two resemble Piccard (1961) §128938 and §71625 respectively.



BnF fr. 1881

BnF fr. 1881, a majority paper manuscript dating to the turn of the sixteenth century, contains some 14 items, but only 7 which are any longer than a couple of folios, with the manuscript's other short texts containing a mixture of poetry and prose in both French and Latin. The longer texts in the manuscript can be split into two categories, namely prose treatises on moral behaviour on the one hand, and saints' lives and miracles on the other. These two categories are mixed within the codex, with no clear division into sections, in such a way that the narrative texts reinforce the moral messages of the treatises, and vice versa.

In the first category, we find Robert le Chartreux's *Le chastel perilleux* (ff. 1r–92v), a French translation of Hugh of Saint Victor's *Tractatus de anima* (ff. 221r–236v), a treatise on Christian perfection of which the beginning is missing (ff. 238r–252r), and *Le miroir des dames mariees* (ff. 286r–308r).[19] Alongside *Robert* (ff. 253r–285v), the second category contains a version of the *Vie de saint Alexis* in alexandrine monorhyme quatrains (ff. 93r–123r), as well as a prose selection of *miracles* of the Virgin Mary (ff. 125r–216r).[20] *Robert* shares a good deal of similarities with the *Vie de saint Alexis* both in content and form, telling of an individual who adopts a mendicant lifestyle as an act of piety. It also tells a comparable story of salvation to those found within

[19] Only the first and last of these texts are edited, in Brisson (1974) and Pichon (1846, i: 99–125) respectively.
[20] The version of the *Alexis* is the so-called *A4* version rejected by Munk Olsen (1963b) as a candidate for inclusion amongst Jehan's *œuvre*, and edited in Paris & Pannier (1872).

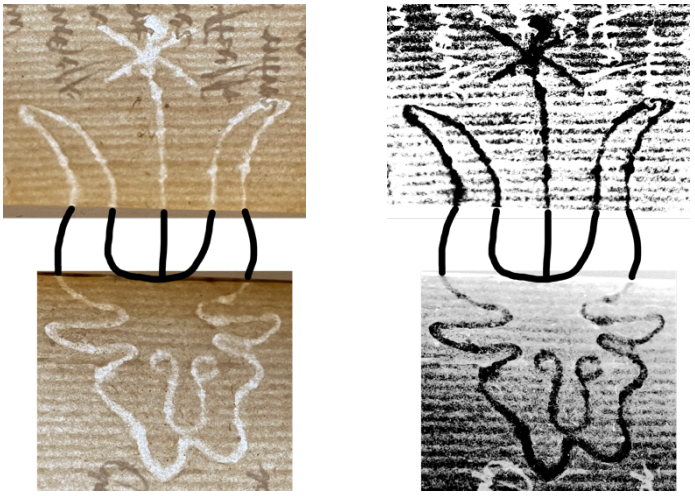
the *miracles*, but it is notable that the compilation of this manuscript appears not to have been carried out so as to gather similar texts together in coherent groupings, since *Robert* is found at a great distance from the other texts of the second category. Rather, given that *Robert* follows the treatise on Christian perfection, it would appear that the text was included to offer narrative evidence for the arguments made within that text, as a story telling of an individual who loses his way but eventually reaches perfect piety and is redeemed.

The rubric of *Robert*, which takes up the whole of f. 253^{r}, also perhaps alludes to a possible reason for the text's inclusion, in that it is described as:

> Ung moult beau livre, lequel parle de la vie d'ung seigneur qui fut nommé Robert le Dyable, lequel fut filz du duc de Normandie et de la fille de Mon*seigneur* le duc de Bourgougne, qui est une belle chose a ouyr.

The increased focus on the fact that Robert's mother is the daughter of a duke of Burgundy, a fact mentioned in *Robert* but which has little bearing on the narrative, suggests possible Burgundian origins for BnF fr. 1881. It also suggests that the compiler-scribe of this section of the codex, likely also a proud Burgundian, included *Robert* in the manuscript primarily because it reinforces the moral messages presented elsewhere, but also perhaps because it allows the compiler-scribe to foreground his own allegiance to the duke of Burgundy.

Figure 3.88 – watermark found in BnF fr. 1881, combined from ff. 257 and 266 with an impression of the missing section.



Again, examination of the manuscript's material composition can be fruitful here. In short, BnF fr. 1881 is somewhat chaotic in its materiality. The codex includes four leaves of parchment, and its paper leaves display some 8 distinct watermarks (split across the spine of the volume given the quarto format), many of which appear at a number of distinct moments in the codex. It is also written in at least four different hands, with shifts on ff. 124, 237, and 238. *Robert* is copied predominantly on paper bearing a watermark depicting a bull's head with a star above linked with a single-contour bar, as displayed in Figure 3.88,[21] as is *Le miroir des dames mariees* and the first 33 folios of *Le chastel perilleux*, the only other texts copied on such paper. Both are texts of the first category, on moral instruction, and so if they were copied at the same time as *Robert* then it would appear that little attempt was made to distinguish between texts of the two different types. The manuscript therefore bears all the hallmarks of one that was compiled over a long period of time, as the compiler-scribe(s) added texts of interest to the codex as they were encountered, using whatever writing materials were accessible at the time.

Besançon Bib. mun. 588

Besançon Bib. mun. 588, the fourth manuscript to contain *Robert*, is once again much later than fr24432, and is indeed the latest manuscript to contain any of the texts which we have attributed to Jehan, dating to the early sixteenth century. It is a paper manuscript, consisting of 172 leaves, and contains many comparable texts to BnF fr. 1881, but unlike the latter codex it contains clearer groups of texts, indicating a greater degree of compilatory intent.

The first group of texts in Besançon Bib. mun. 588, which we may consider to function as a sequence, is concerned with the Passion, the events of which are retold in prose (ff. 1^{r}–25^{r}) before the inclusion of the prose version III of *La vengeance de Notre Seigneur* (ff. 25^{v}–41^{r}), a text

[21] Resembling Piccard (1961) §80274.

which tells the apocryphal story of the destruction of Jerusalem in the years following the death of Christ.[22] This text is introduced in the manuscript as the 'Vengeance de la mort et passion de nostre Seigneur Jhésu Crist et aussi de la destruccion de Jherusalem et des Juis,' highlighting its themes of righteous retribution, and the link with the preceding text.

This sense of divine justice is reinforced in the second major sequence within the codex, but this time less from the perspective of divine judgement, and rather from that of divine salvation. The group begins with a prose text, *Le remede de l'ame* (ff. 41^{v}–54^{v}),[23] a penitential text containing information for sinners about how to redeem their souls. To reinforce the messages of this text, the compiler then includes *Robert* (ff. 55^{r}–63^{v}) and, just as in BnF fr. 1881, a version of the *Vie de saint Alexis* in alexandrine monorhyme quatrains (ff. 64^{r}–73^{v}). Thus, while the *Alexis* does not appear in fr24432, it was clearly a text with which *Robert* was heavily associated, given the co-occurrence of these two texts in two of the manuscripts containing Jehan, most likely on account of their narrative and didactic similarities.

A further text which would also appear to fit with this group is Jehan le Fèvre de Ressons' translation of the *Disticha Catonis* (ff. 85^{r}–92^{v}), a collection of proverbial wisdom whose content reinforces the didactic content of this second sequence.[24] That said, the extent to which this text can be seen to belong to the sequence concerning divine salvation is limited by the intervening appearance of a series of 'armoines et ordonnances' (ff. 74^{r}–84^{v}). These line up much more closely with the final grouping of the manuscript which contains letters, agreements, and edicts concerned primarily with the administration of Besançon and its surrounding area, reinforcing the suggestion that the manuscript was originally compiled in the town in which it now resides. Interestingly enough, despite the fact that the manuscript

[22] For an edition, see Graf (1882–1883).
[23] The text appears to be unedited.
[24] For details of editions, see note 25.

dates from the early sixteenth century, most of these administrative texts can be straightforwardly dated to within the fifteenth-century, with the earliest such text being an agreement between the commune of Besançon and its clergy (ff. 126r–132v), which dates to 1435. The compiler of Besançon Bib. mun. 588 therefore appears to have chosen texts with a significance that extends beyond his immediate period. It is therefore also possible that such a motivation applied to his inclusion of *Robert* and the alexandrine *Alexis*, since both these versions of their respective tales had likely been in circulation for some two centuries before their appearance in this particular codex.

Across the four manuscripts in which it appears, then, *Robert* is included as an example in support of moral messages found elsewhere. It is also, through its co-occurrence with both the alexandrine *Alexis* and the *Vie des peres*, presented alongside some of the best known tales of saints' lives, suggesting a hagiographic interpretation of the text which is not quite as prominent in fr24432.

3D.3.3 Florence *in Grenoble Bib. mun. 871*

Grenoble Bib. mun. 871 is a composite manuscript dating to the fifteenth century, which consists of three separate booklets. The first (ff. 1r–107v) appears to have been conceived as a collection of proverbs, sayings, and teachings, since all of its texts consist entirely or in part of such instructions. It begins with Christine de Pizan's *Proverbes moraux* (ff. 1r–4v), followed by *Les dits et proverbes des sages* (ff. 5r–13v), a vernacular translation of the *Liber floretus* (ff. 14r–56r), George Chastelain's *Le miroir de mort* (ff. 56v–71v), Jehan le Fèvre de Ressons' translation of the *Disticha Catonis* (ff. 72r–98v, the same text discussed above in the context of *Robert*) and finally

Version I of the *Facet en françoys* (ff. 99r–107r), another vernacular translation of a Latin text, this time the *Facetus*.[25]

The second booklet of the codex (ff. 108r–187v), with the exception of a paschal table on f. 187v, is taken up by Jacques le Grand's *Livre des bonnes mœurs* (ff. 108r–187v),[26] and was perhaps therefore included to add an extra text considered relevant to the contents of the first section.

In the third and final section (ff. 188r–211v), the focus of the codex shifts, since this section appears to have been conceived as a collection of the lives of holy individuals. *Florence* (ff. 188r–199r) begins the section, being introduced as a saint's life using the formula 'La vie Fleurence de Rome' in the manuscript's rubrication. It is followed by *La patience de Griseldis* (ff. 199r–206r),[27] whose eponymous main character closely resembles Florence, being a young noblewoman who undergoes various trials before regaining her rightful place beside her husband. The text is in prose, and therefore the juxtaposition has almost certainly been made exclusively on the basis of the thematic similarities which these two texts share, and the focus which this allows on the texts' moral messages, particularly concerning female behaviour. Indeed, the compiler-scribe of this section of the manuscript appears to acknowledge this, writing after the explicit to *Florence* on 199r that:

> A l'exemple des fames mariés ay cy mis ung petit livret appelé le Mirouer des Dames, d'une vaillent fame nommee Griselidis et du marquis de Saulces, son mari.[28]

Following a similar moralising section by the compiler-scribe at the end of *Griseldis*, we find *La Vie de saint Jehan Paulus* (ff. 207r–209r). Being a text about a male saint whose trials are of a different nature from those of Florence or Griseldis, Jehan Paulus' *vie* does not match the

[25] For editions of all except the translation of the *Liber floretus*, which remains unedited in its entirety, see Kosta-Théfaine (2010), Morawski (1924), van Hemelryck (1995), Ulrich (1904), and Morawski (1923) respectively. The three Latin texts are edited in Orbán (1978), Chase (1922), and Morawski (1923) respectively.
[26] For an edition, see Beltran (1986).
[27] For an edition, see Nodier (1843).
[28] The transcription is that of Walton (1928).

preceding texts quite so closely in its content, but it does at least share some formal similarity with *Florence* in that it is written in alexandrine monorhyme quatrains.[29] Finally, a possible desire for numeric balance between female and male saints is indicated by the presence of the last text in the text, a prose tale, the *Vie de saint Georges, martyr* (ff. 209r–211v).[30]

In Grenoble Bib. mun. 871, then, *Florence* is presented very differently from in fr24432, which would have had a significant impact on how a reader would have encountered the text. It is no less didactic in this codex, as evidenced by the compiler-scribe's additions, but while the universal, penitential content of the tale would have come to the fore as part of the 'long-text' grouping in fr24432, the compilation in Grenoble Bib. mun. 871 points towards a reading which is far more about the adoration and imitation of a particular individual. In this reading, *Florence* becomes a celebration of a *mulier sancta*, serving as a model for pious married women, rather than a general example used as an invitation to all readers to change their ways, as it is in fr24432.

3D.3.4 Anelés *in BnF fr. 5036*

BnF fr. 5036, despite being a paper manuscript, largely written in single columns, which dates from the second half of the fifteenth century, is in many ways comparable to fr24432, in that it contains a number of different texts and text types, in both poetry and prose, often with no clear indication of the compilatory motivations behind their inclusion.

Flanked at either end by treaties and other diplomatic correspondence (ff. 2r–40v and 300r–322r), the main body of texts contained within the codex predominantly consists of a mixture

[29] The text is absent from the major directories of medieval narrative poems, including Naetebus (1891), and is unedited, but the *Nouveau Naetebus* (Seláf 2011: http://nouveaunaetebus.elte.hu/adatlap.php?m_index=49) lists it as being written in such a form. For an extended explicit, see Walton (1928: 474).
[30] Both texts remain unedited.

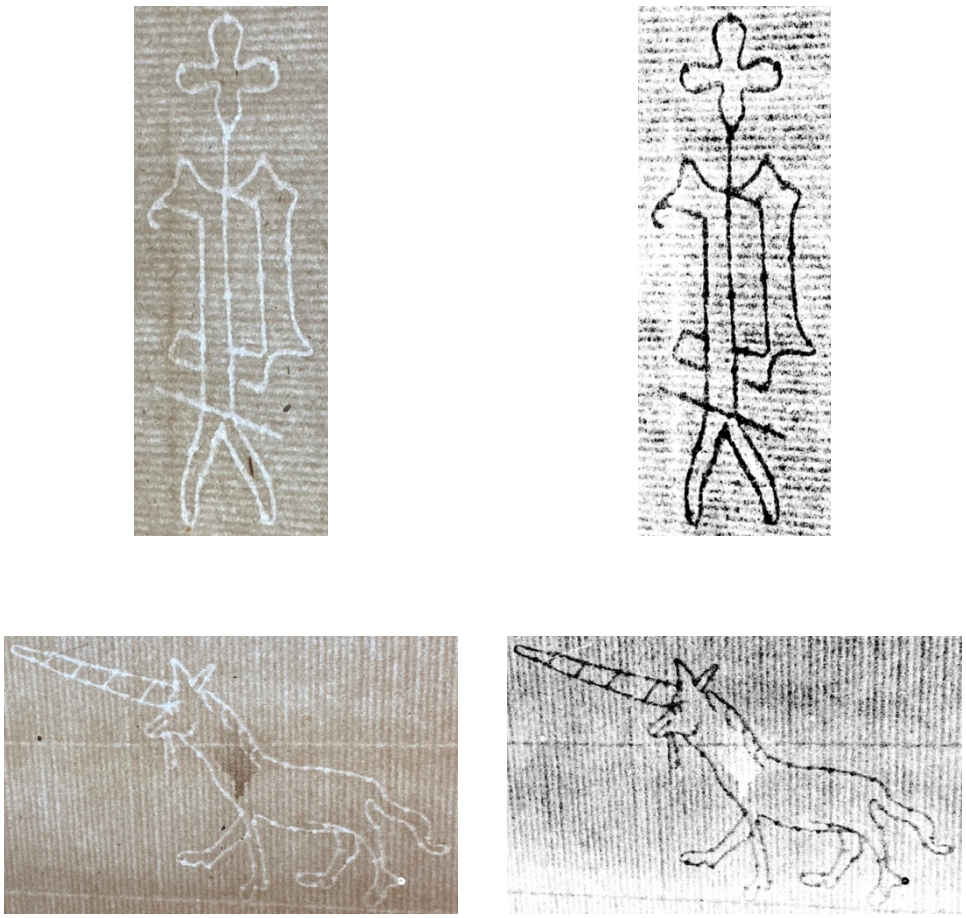
of praise of the Blessed Virgin (e.g. ff. 197r–215r, 230r–235r) and moralising stories (e.g. ff. 42r–125r, 236r–299v). *Anelés* falls into this second category, appearing on ff. 216r–225v.

Where the content of BnF fr. 5036 differs from fr24432, however, is in the fact that the former codex is bilingual, containing texts in both French and Latin, with the compiler seemingly unconcerned by such code-switching. The compiler of fr24432, by contrast, appears to have consciously decided to compile the manuscript almost exclusively in French, offering vernacular translations of texts otherwise widely transmitted in Latin, which themselves include Latin only when absolutely necessary, for example to refer to elements of the Latin mass.[31]

Notably, the copy of *Anelés* in BnF fr. 5036 appears between two Latin prose texts, an account of the lamentations of Mary entitled the *Compassio beate Marie circa crucem* on the one side (ff. 209r–215r), and an extended General Confession, entitled *Exemologesis, seu confessio generalis peccatorum valde pia, cum aliquot devotis orationibus* (ff. 226r–228v) on the other.[32] It is this latter text which appears most obviously connected to *Anelés*, since the story's primary didactic function is the demonstration of the importance of confession, as a tale about how a family, tainted by incest, is raised to sainthood through confession and appropriate penance. Given that *Anelés* is copied on paper bearing the same unicorn watermark as that found for the *Compassio*, rather than the gothic 'p' with a four-leaf flower found for the *Exemologesis* (see Figure 3.89),[33] *Anelés* appears to have been consciously compiled as a precursor to the *Exemologesis* at some point after the texts were copied, in order to convince a reader of the importance of confession before the later presentation of the general confession.

[31] Such an example is found in *Miroir*, a vernacular translation of the *Speculum ecclesiae* (ff. 317rb–327rb).
[32] Neither text appears to have been edited.
[33] Resembling Piccard (1961) §124163 and §113248 respectively.

Figure 3.89 – watermarks found in BnF fr. 5036 (top: f. 14, bottom: f. 7).



In this sense, while its two linguistic and textual contexts are distinct across both manuscripts, the compilation of *Anelés* in BnF fr. 5036 appears somewhat comparable to that in fr24432, since the text's didactic function is foregrounded in both cases. In the former codex, this is due to the inclusion of *Anelés* alongside a more explicitly liturgical text which nevertheless shares a focus on confession, while in the latter, it is due to its inclusion in the 'long-text' grouping, whereby it is placed alongside other long texts with a similarly clear didactic focus.

3D.3.5 Sauveur *in BnF Ars. 2115*

BnF Ars. 2115, the manuscript containing the only known copy of *Sauveur*, is a paper codex compiled towards the end of the fifteenth century. It consists of some 67 leaves but may once

have been intended to be longer, since the final text is incomplete.[34] Its compilatory motivations are made clear by a notice on f. 5, which reads:

> Ce qui est en ce petit livre doibvent enseigner les prestes a leurs paroissiens, et pour les simples prestes qui n'entendent pas bien l'Escripture, et pour les simples gens, et fait en franchois par grant conseil […][35]

Unlike BnF fr. 5036, then, BnF Ars. 2115 is apparently conceived with the conscious intention of offering moral lessons exclusively in French, for those who might otherwise be unable to access texts written in Latin. In this sense, it is certainly comparable to fr24432, and these aims do apparently coincide with the texts contained within the codex, which appear on the whole to have some didactic or parenetic function, including *Le dist des VII pechés mortelx* (f. 35^{v}–38^{r}), and *Des dix des sages* (f. 59^{v}–65^{r}).

Sauveur is split within the manuscript, due to the later erroneous insertion of three sheets in the same hand from elsewhere in the same manuscript, with the first part spanning the recto and verso of f. 48, and the remainder of the text appearing on ff. 52^{r}–57^{v}.[36] The text which follows *Sauveur* (f. 58^{r}–59^{r}), entitled *Une exemple de tribulacion*, is a very short poem foregrounding the faith of a woman who has suffered in this life but takes comfort from the prospect of the next. The content of this poem, combined with its amusingly non-specific title, therefore reinforces the explanation for the inclusion of *Sauveur* in this section of the manuscript, since *Une exemple de tribulacion* could just as well be used to describe the plot of Jehan's text, which tells the tale of a young man, whose mother promised him to the Devil, who saves himself through a life of seclusion. In BnF Ars. 2115, then, it is both *Sauveur*'s choice of language and its content which

[34] This incompleteness does not result from the later loss of material, however, since the table of contents ends at the same point, and the verso side of the final leaf is blank.
[35] The transcription is that given by Martin (1885–1899, ii: 420).
[36] The error is likely to be that of a later individual to the scribe of this section, and can therefore be ignored.

mark it as an acceptable candidate for inclusion in the manuscript, meeting both of the stated aims of the compiler.

3D.3.6 Abesse *in BnF fr. 12483*

The final text attributed to Jehan which we are yet to consider is *Abesse*, which appears only within BnF fr. 12483, the so-called *Rosarius* manuscript, which is a parchment manuscript generally accepted to date to the second quarter of the fourteenth century, that is to say within a decade or so of the compilation of fr24432, itself compiled towards the very middle of the century. Much has been written on the *Rosarius*,[37] but suffice to say that rarely is a manuscript so clear in its compilatory intent: the manuscript's content is dedicated in its entirety to praise and honour of the Blessed Virgin, and the compiler-author-scribe, likely a member of the Dominican order at Poissy (Långfors 1916: 515), is meticulous in the format and structure followed by the 50 chapters in each of the manuscript's two volumes.[38]

The structure for each of the surviving chapters is largely the same, as Långfors explains (1916: 504):

> Chacun est précédé d'une rubrique et d'un numéro d'ordre. En tête se trouve la description, d'une trentaine de vers en moyenne, d'un animal, d'une pierre, d'une plante, d'une « chose » quelconque (le mot est de l'auteur). Ces faits plus ou moins exacts, empruntés à l'histoire naturelle, servent de fondement à un enseignement moral. Car, après avoir énoncé plus ou moins longuement les différentes propriétés d'une « chose », l'auteur les reprend une à une pour les rapprocher des qualités similaires de la Vierge, dont les vertus se trouvent ainsi symbolisées autant de fois.

The chapter then continues with one or two devotional tales, most frequently Marian miracles, and then closes with a song, *lai*, or *dit*, tied by the compiler-author-scribe into the broader structure of the chapter.

---

[37] A comprehensive study of the manuscript is to be found in Långfors (1916). Other, more recent, examination includes Kunstmann (1991), Zetterberg and Sandqvist (1994), Mattiacci (1996), and Sandqvist (1996).
[38] That is, as far as we can tell, since only 57 of the original 100 chapters survive (Mattiacci: 1996: 4), and many of the remaining folios have been badly mutilated.

Abesse appears on ff. 66ra–68rb, in chapter 36. The *chose* for this chapter is parsley, and the intention of the chapter appears to be a presentation of how Mary helps women whose pregnancies continue into post-term. The poem provides the culmination to the chapter, and the reasons for its inclusion are made crystal clear by the compiler-author-scribe in the lines immediately preceding the text:

> Du confort la douce Marie,
> Co*m*ment a une sienne amie
> Aida jadis a son besoing,
> Un tel dit en ay a tesmoing.
> Q*ui* le dist fist je ne sai mie,
> Mes ainsi p*ar*le de Marie. (f. 66ra)

Abesse was thus included in the *Rosarius* almost solely on thematic grounds, offering, as it does, a story about an abbess' salvation at the hands of Mary after falling pregnant by an abbot. The text does not refer to post-term pregnancy, but the compiler-author-scribe clearly considered it a sufficiently good example of the assistance that Mary offers to pregnant women, that he felt that it warranted inclusion in this section.

While the text's content is undoubtedly the most important reason for its inclusion, we should note that both its language and form appear to have been deciding factors, too. Since the *Rosarius* is primarily written in French, it is possible that *Abesse* was chosen ahead of its Latin analogues on account of its choice of language.[39] The fact that it is written in verse, rather than prose, also appears to have played a role, since this appears to be a requirement for texts at the end of chapters in the *Rosarius*, but the exact nature of this verse appears not to have been important, since the surrounding text, including that which introduces *Abesse*, is octosyllabic.

[39] On these analogues, see Munk Olsen (1978: cxiii–cxvii).

The compiler-author-scribe's assertion that 'Qui le dist fist je ne sai mie' is tantalising, revealing his interest in the authorship of the poems he uses, but acknowledging that such information was not forthcoming for *Abesse*. This interest in authorship is observed elsewhere in the *Rosarius*, where specific claims are made as to who wrote a given text, and also in the use of marginal notes (usually an abbreviated form of *Quidem*) which indicate that a text stems from an external source. In its contrast with the expression of the poem's content, 'Mes ainsi parle de Marie,' this phrase also indicates that the compiler-author-scribe's interest in authorship did not necessarily extend to a compilation based on grounds of authorship, rather a text's themes were considered sufficient reason for its inclusion in a given chapter.

That the *Rosarius* dates from a similar period to fr24432 therefore supports the hypothesis that the compiler of the latter manuscript may have been unaware of any shared authorship between the texts now attributed to Jehan. At least *Abesse* appears to have been circulating with no indication of who wrote it, with even a compiler interested in authorial attribution unable to find out such information, and so it is possible that the same is true for the rest of Jehan's corpus, particularly those texts that appear within fr24432. This would therefore help to explain why the texts are grouped primarily on thematic and formal grounds, while authorship does appear to be important elsewhere in the codex, for example through the grouping of Rutebeuf's texts.

The analysis in this section has displayed the value in taking a micro-level codicological approach to the manuscripts which contain the poems which we have attributed to Jehan. Such an approach can help us gain a clearer understanding of the elements of the texts which were considered important by the manuscripts' compilers, and therefore perhaps also by the texts' early readers. Questions of authorship do not appear to have been at the forefront of all these compilers' minds, but some do nevertheless appear to have been preoccupied by such questions. Instead, the compilatory procedures appear first and foremost to have been

motivated by the formal, thematic, and linguistic features of the texts. These, we may note, are the same features which we have considered throughout Chapter 3 in reinforcing the claim that these texts are ultimately the work of a single individual.

Acknowledgements

First and most significant thanks must of course go to my supervisor, Prof Daron Burrows, for being a constant font of knowledge and advice, be it academic, pedagogical, or professional.

I am grateful also to the rest of the Oxford French community, notably Profs Helen Swift and Sophie Marnette, whose feedback at the various milestones of the course has been invaluable, and Dr Luke O'Sullivan, my college advisor. Thanks go also to my fellow DPhil candidates, particularly Jack Nunn and Dr Anna Wilmore, who have offered insight from their own specialisms, and the occasional act of libatory solidarity.

For teaching me to code, I am indebted to Prof Henrike Lähnemann, Emma Huber, Prof John Chelsom, Steven Pemberton, and Dr Charles Redmon, as I am to the staff of the Bodleian Libraries, the British Library, the Bibliothèque nationale de France, and the library of Merton College, in particular Dr Laure Miolo, Dr Alexandre Tur, Jérémy Chaponneau, and Dr Julia Walworth, for facilitating my encounters with the manuscripts in their collections. On the topic of manuscripts, I am grateful to my co-organisers at Teaching the Codex, particularly Drs Mary Boyle and Tristan Franklinos, for the opportunities to remind myself how exciting manuscript studies is.

The methodology presented in Chapter 2 is the product of airings at the 2022 TEI Conference, the 2023 International Congress on Medieval Studies, and the 2023 and 2024 International Medieval Congress, and so I am grateful to the attendees of these conferences, as well as to Profs Richard Trachsler, Baudouin van den Abeele, and Jonathan Prag, and Drs Henry Ravenhall, Luca Barbieri, and Robert Crellin for their indispensable feedback.

I am grateful to Dr Katherine Lunn-Rockliffe, Prof Ian Maclachlan, and the teaching staff and students at Hertford, Merton, and St Peter's colleges, who have broadened my horizons

and kept me grounded in French as a living language. Thanks also go to Amy Thompson, for reminding me what busy looks like, and to Drs Clement Collins-Rice, Thomas Miller, and Joe Morford, for providing me with the epistemic humility which we humanists can only reach through encounters with the sciences.

Finally, none of the above would have been possible without money, and so I am grateful to Prof Jeanette Beer and Sir Basil Zaharoff (†) for the funding for my studentship, and to Prof David Wulstan (†) for providing funding for archival and conference trips, and for digitisation costs.