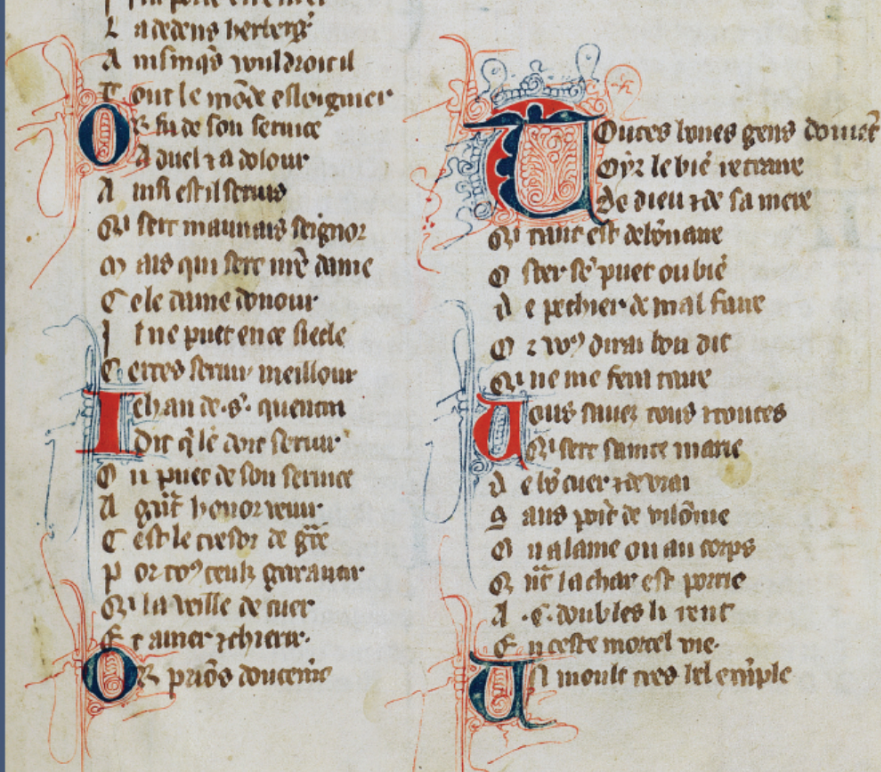




# MEDIEVAL AUTHORSHIP AND STATISTICAL ANALYSIS

## RECONSIDERING JEHAN DE SAINT-QUENTIN'S ŒUVRE IN THE DIGITAL AGE

by

SEBASTIAN DOWS-MILLER

ARC HUMANITIES PRESS



## **BOOK CULTURES, MEDIEVAL TO MODERN**

### **Further Information and Publications**

[www.arc-humanities.org/series/book-series/](http://www.arc-humanities.org/series/book-series/)

# **MEDIEVAL AUTHORSHIP AND STATISTICAL ANALYSIS**

## **RECONSIDERING JEHAN DE SAINT-QUENTIN'S ŒUVRE IN THE DIGITAL AGE**

by

**SEBASTIAN DOWS-MILLER**

**ARC**HUMANITIES PRESS

## British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library.

© 2026, Arc Humanities Press, Leeds



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Licence.

The author(s) assert(s) their moral right to be identified as the author(s) of their part of this work.

Permission to use brief excerpts from this work in scholarly and educational works is hereby granted provided that the source is acknowledged. Any use of material in this work that is an exception or limitation covered by Article 5 of the European Union's Copyright Directive (2001/29/EC) or would be determined to be "fair use" under Section 107 of the U.S. Copyright Act September 2010 Page 2 or that satisfies the conditions specified in Section 108 of the U.S. Copyright Act (17 USC §108, as revised by P.L. 94-553) does not require the Publisher's permission.

ISBN (Hardback): 9781802703269

ISBN (Paperback): 9781802704853

e-ISBN (PDF): 9781802704839

e-ISBN (ePUB): 9781802704846

DOI: 10.17302/ogkl5486

**www.arc-humanities.org**

Printed and bound in the UK (by CPIGroup [UK] Ltd), USA (by Bookmasters), and elsewhere using print-on-demand technology.

**Publisher (manufacturer) details:** Arc Humanities Press, 14 Clifton Moor Business Village, James Nicolson Link, York YO30 4XG, United Kingdom.

**EU Authorized Representative details (for GPSR purposes):** Amsterdam University Press, Nieuwe Prinsengracht 89, 1018 VR Amsterdam, The Netherlands. [www.aup.nl](http://www.aup.nl)

## CONTENTS

|                                                  |     |
|--------------------------------------------------|-----|
| List of Illustrations.....                       | vi  |
| Acknowledgements.....                            | x   |
| Abbreviations.....                               | xi  |
| Introduction .....                               | 1   |
| Chapter 1. Methodology.....                      | 17  |
| Chapter 2. Form.....                             | 29  |
| Chapter 3. Genre.....                            | 39  |
| Chapter 4. Terms of Address.....                 | 61  |
| Chapter 5. Possessive Adjectives.....            | 77  |
| Chapter 6. Versification.....                    | 87  |
| Chapter 7. Shared Vocabulary and Structures..... | 115 |
| Chapter 8. Sources.....                          | 141 |
| Conclusion.....                                  | 159 |
| Bibliography .....                               | 163 |
| <i>Index operum</i> .....                        | 171 |

# LIST OF ILLUSTRATIONS

## Figures

|                                                                                                                                                                  |     |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 1. Custom R function to calculate a p-value for the use of a given feature between a target and reference corpus using randomized sampling. ....          | 21  |
| Figure 2. Box-and-whisker plot of the lengths of the <i>J-corpus</i> .....                                                                                       | 31  |
| Figure 3. Box-and-whisker plot of the lengths of the texts of the <i>J-corpus</i> , compared to those of the <i>P1-corpus</i> . ....                             | 32  |
| Figure 4. Box-and-whisker plot of the lengths of the texts of the <i>J-corpus</i> , compared to those of the <i>A</i> -, <i>F</i> -, and <i>R-corpora</i> . .... | 32  |
| Figure 5. Rates of use (x-axis) and raw numbers (figures) of stanzas containing feminine rhymes for each text of the <i>J-corpus</i> . ....                      | 90  |
| Figure 6. Rates of occurrence for runs of stanzas containing masculine and feminine rhymes across the <i>J-corpus</i> .....                                      | 91  |
| Figure 7. Box and whisker plots for the rate of use of feminine rhymes in the <i>J</i> - and <i>A-corpora</i> .....                                              | 93  |
| Figure 8. Rates of use (x-axis) and raw numbers (figures) of stanzas containing feminine rhymes for each text of the <i>A-corpus</i> .....                       | 94  |
| Figure 9. Rates of occurrence for runs of stanzas containing masculine and feminine rhymes in the <i>A-corpus</i> . ....                                         | 95  |
| Figure 10. Rates of use (x-axis) and raw numbers (figures) of stanzas containing “easy” rhymes in the <i>J-corpus</i> .....                                      | 98  |
| Figure 11. Box and whisker plots for the rate of use of “easy” rhymes in the <i>J</i> - and <i>A-corpora</i> .....                                               | 99  |
| Figure 12. Rates of use (x-axis) and raw numbers (figures) of stanzas containing “easy” rhymes in the <i>A-corpus</i> . ....                                     | 100 |
| Figure 13. Box and whisker plots for the rate of use of “easy” rhymes in the <i>J</i> -, <i>A</i> -, and <i>R-corpora</i> . ....                                 | 101 |

|                                                                                                                                                |     |
|------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 14. Scatter plot for the rhyme coefficients of lemmata<br>found in rhyming position in the <i>J-</i> and <i>A-corpora</i> . ....        | 108 |
| Figure 15. Rates of use of each type of discourse onset in the <i>J-corpus</i> . ...                                                           | 133 |
| Figure 16. Rates of use of each type of discourse onset in the texts<br>of the <i>A-corpus</i> that contain instances of direct discourse..... | 136 |

## Tables

|                                                                                                                    |    |
|--------------------------------------------------------------------------------------------------------------------|----|
| Table 1. Texts attributed to Jehan de Saint-Quentin by Gröber<br>(1888–1902).....                                  | 7  |
| Table 2. Texts attributed to Jehan de Saint-Quentin by Morawski<br>(1939, 328–30).....                             | 7  |
| Table 3. Texts attributed to Jehan de Saint-Quentin by Munk Olsen (1978)....                                       | 9  |
| Table 4. Verses 1–4 of <i>Abesse</i> , tagged and lemmatized by RNNTagger.....                                     | 24 |
| Table 5. Verses 1–4 of <i>Abesse</i> , tagged and lemmatized by ChatGPT. ....                                      | 24 |
| Table 6. Uses of <i>miracle</i> as a generic descriptor within the <i>J-corpus</i> .....                           | 43 |
| Table 7. The single use of <i>miracle</i> as a generic descriptor within<br>the <i>A-corpus</i> . ....             | 43 |
| Table 8. Other uses of <i>miracle</i> alongside <i>bel</i> within the <i>J-corpus</i> .....                        | 44 |
| Table 9. Uses of <i>miracle</i> alongside <i>bel</i> within the <i>A-corpus</i> .....                              | 44 |
| Table 10. Uses of additional descriptors alongside <i>miracle</i> within<br>the <i>A-corpus</i> . ....             | 45 |
| Table 11. Uses of <i>miracle</i> alongside verbs of speaking and hearing<br>within the <i>J-corpus</i> . ....      | 46 |
| Table 12. Uses of <i>exemple</i> and <i>exemplaire</i> as generic descriptors<br>within the <i>J-corpus</i> . .... | 50 |
| Table 13. Uses of <i>exemple</i> as a generic descriptor within the <i>A-corpus</i> .....                          | 51 |
| Table 14. Uses of <i>dit</i> as a generic descriptor within the <i>J-corpus</i> . ....                             | 55 |
| Table 15. Uses of <i>dit</i> as a generic descriptor within the <i>A-corpus</i> . ....                             | 56 |

|                                                                                                                                                                    |    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Table 16. Uses of <i>bone gent</i> as a means of referring to the audience within the <i>J-corpus</i> . . . . .                                                    | 63 |
| Table 17. Uses of <i>bone gent</i> as a means of referring to the audience within the <i>A-corpus</i> . . . . .                                                    | 63 |
| Table 18. Uses of <i>bone gent</i> within the <i>J-corpus</i> for purposes other than to address the audience. . . . .                                             | 65 |
| Table 19. Uses of <i>seigneur</i> as a means of referring to the audience within the <i>J-corpus</i> . . . . .                                                     | 68 |
| Table 20. Hierarchical uses of <i>seigneur</i> within the direct discourse of the <i>J-corpus</i> . . . . .                                                        | 68 |
| Table 21. Egalitarian uses of <i>seigneur</i> within the direct discourse of the <i>J-corpus</i> . . . . .                                                         | 69 |
| Table 22. Uses of <i>seigneur</i> as a means of referring to the audience within the <i>A-corpus</i> . . . . .                                                     | 70 |
| Table 24. Uses of <i>tous et toutes</i> and <i>ceulz and celles</i> as means of addressing the audience in the <i>A-corpus</i> . . . . .                           | 72 |
| Table 25. Examples of elided <i>ma</i> etc. preceding feminine vowel-initial nouns within the <i>J-corpus</i> . . . . .                                            | 78 |
| Table 26. Examples of <i>mon</i> etc. preceding masculine vowel-initial nouns within the <i>J-corpus</i> . . . . .                                                 | 78 |
| Table 27. Uses of <i>erreur</i> within the <i>J-corpus</i> . . . . .                                                                                               | 79 |
| Table 28. Examples of <i>mon</i> etc. before unambiguously feminine vowel-initial nouns in the <i>P-corpus</i> . . . . .                                           | 82 |
| Table 29. Uses of <i>mon</i> etc. before vowel-initial feminine nouns in the copy of <i>Florence</i> found in Grenoble, BM, MS 319, compared to fr. 24432. . . . . | 83 |
| Table 30. Examples of <i>mon</i> etc. before unambiguously feminine vowel-initial nouns in <i>Jehan Paulus</i> and <i>Richard</i> . . . . .                        | 84 |

|                                                                                                                                                                                                                                            |     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Table 31. Examples of <i>ma</i> etc. before feminine vowel-initial nouns<br>in <i>Jehan Paulus</i> and <i>Richard</i> .....                                                                                                                | 84  |
| Table 32. Counts and rates of lemmata which appear on ten or<br>more occasions in the <i>J-corpus</i> , limited to the top twenty results.....                                                                                             | 105 |
| Table 33. Counts and rates of lemmata which appear on ten or<br>more occasions in the <i>A-corpus</i> , limited to the top twenty results.....                                                                                             | 106 |
| Table 34. The six most common lemmata at the beginning of a<br>line for each of the texts in the <i>J-corpus</i> .....                                                                                                                     | 121 |
| Table 35. The six most common lemmata at the beginning of a<br>line across the entire <i>J-corpus</i> .....                                                                                                                                | 122 |
| Table 36. Results of p-value tests comparing use of the six most<br>frequent lemmata at the beginning of a line across the entire<br><i>J-corpus</i> .....                                                                                 | 123 |
| Table 37. The six most common lemmata at the beginning of a<br>line across the <i>J</i> -, <i>A</i> -, <i>F</i> -, and <i>R-corpora</i> .....                                                                                              | 124 |
| Table 38. Average rates of the use of the six most common<br>lemmata at the start of lines in the <i>J-corpus</i> , compared to the<br><i>A</i> -, <i>F</i> -, and <i>R-corpora</i> .....                                                  | 125 |
| Table 39. Results of p-value tests comparing use of the six most<br>frequent lemmata at the beginning of a line across the entire<br><i>J-corpus</i> to <i>Guillaume</i> , <i>Robert</i> , <i>Richard</i> , and <i>Jehan Paulus</i> . .... | 126 |
| Table 40. Counts of each type of discourse onset for each text in<br>the <i>J-corpus</i> . ....                                                                                                                                            | 131 |
| Table 41. Uses of the phrase <i>trouver lisant</i> in the <i>J-corpus</i> . ....                                                                                                                                                           | 145 |
| Table 42. Uses of the phrase <i>trouver lisant</i> in the <i>A-corpus</i> . ....                                                                                                                                                           | 145 |
| Table 43. References to written sources in the texts of the<br><i>J-corpus</i> , as well as <i>Guillaume</i> and <i>Robert</i> . ....                                                                                                      | 147 |

## ACKNOWLEDGEMENTS

MUCH OF THE content of this book arose out of my doctoral thesis. First and most significant thanks go therefore to my DPhil. supervisor, Prof. Daron Burrows, for being a constant font of knowledge and advice, be it academic, pedagogical, or professional. I am grateful also to my examiners, Professors Sophie Marnette, Helen Swift, and Richard Trachsler, for their invaluable feedback at the various milestones of the degree, and their suggestions on how to improve my work for publication. Thanks go also to my fellow DPhil. candidates, particularly Jack Nunn and Dr. Anna Wilmore, who 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 British Library and the Bibliothèque nationale de France, for facilitating my encounters with the manuscripts in their collections. On the topic of manuscripts, I am grateful to Drs. Mary Boyle, Tristan Franklino, Laure Miolo, and Julia Walworth for the opportunities to remind myself how exciting manuscript studies is. Thanks also go to Dr. Katherine Lunn-Rockcliffe, Prof. Ian MacLachlan, and the teaching staff and students at Hertford, Merton, and St. Peter's Colleges in Oxford, who broadened my horizons over the course of this project, and kept me grounded in French as a living language. I am similarly grateful to Professors Jane Gilbert, Jennifer Rushworth, Claire Thomson, and many others at University College London for welcoming me so warmly to my new post and for giving me the space to put the finishing touches to this book. I also gratefully acknowledge UCL's support in providing funding to make the book available in Open Access, and that of the Leverhulme Trust for my new project. Additional thanks 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. 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 DPhil. studentship, and to Prof. David Wulstan (†) for providing funding for archival and conference trips, and digitization costs. Finally, I dedicate this book to the three of my grandparents who saw me start this project, but not finish it: Trevor Miller, Maureen Miller, and Mike Dows. *Diex qui te racheta, ne t'a pas oublié* (Buef, v. 140).

## ABBREVIATIONS

### Lists of Short Titles

#### *J*-corpus

| Full title in Munk Olsen (1978)                                  | Short title                 |
|------------------------------------------------------------------|-----------------------------|
| <i>Un dist que on clamme respon</i>                              | <i>Respon</i>               |
| <i>Le dit des trois chanoïnes</i>                                | <i>Chanoïnes</i>            |
| <i>Le dit des III pommes</i>                                     | <i>Pommes</i>               |
| <i>Le dit de la bourjoise de Romme</i>                           | <i>Romme</i>                |
| <i>Le dit des II chevaliers</i>                                  | <i>Deux chevaliers</i>      |
| <i>Le dit de l'enfant rosti</i>                                  | <i>Enfant rosti</i>         |
| <i>Le dit du povre chevalier</i>                                 | <i>Povre chevalier</i>      |
| <i>Du chevalier et de l'escuier</i>                              | <i>Chevalier et escuier</i> |
| <i>Le dit de la borjoise de Narbonne</i>                         | <i>Narbonne</i>             |
| <i>Le dit du chevalier qui devint hermite</i>                    | <i>Chevalier hermite</i>    |
| <i>Le dit du cordouanier</i>                                     | <i>Cordouanier</i>          |
| <i>Le dit du petit Juitel</i>                                    | <i>Juitel</i>               |
| <i>Le dit de l'enfant qui sauva sa mere</i>                      | <i>Enfant sauva mere</i>    |
| <i>Le dit de l'eau beneoite et du vergier</i>                    | <i>Eaue et vergier</i>      |
| <i>Le dit du riche home qui geta le pain a la teste du povre</i> | <i>Pain</i>                 |
| <i>Le dit du chien et du mescreant</i>                           | <i>Mescreant</i>            |
| <i>Le dit de la pecherresse qui estrangla III enfans</i>         | <i>Pecherresse</i>          |
| <i>Le dit de Merlin Mellot</i>                                   | <i>Merlin</i>               |
| <i>De Florence de Romme</i>                                      | <i>Florence</i>             |
| <i>Le dit des anelés</i>                                         | <i>Anelés</i>               |
| <i>Le dit du buef</i>                                            | <i>Buef</i>                 |
| <i>Le dit de la beguine</i>                                      | <i>Beguine</i>              |
| <i>La Vie saint Sauveur l'ermite</i>                             | <i>Sauveur</i>              |
| <i>D'une abesse que Nostre Dame delivra de confusion</i>         | <i>Abesse</i>               |

A-corpus

| Full title                                       | Short title               | Edition(s) consulted                            |
|--------------------------------------------------|---------------------------|-------------------------------------------------|
| <i>La vie de saint Alexis</i>                    | <i>Alexis</i>             | Paris and Pannier (1872, 327–88)                |
| <i>La vie sainte Barbe</i>                       | <i>Barbe</i>              | Kleist (1973, 277–87)                           |
| <i>La vie saint Basile</i>                       | <i>Basile</i>             | Denomy (1956)                                   |
| <i>La vie de monsieur saint Cristofle</i>        | <i>Cristofle</i>          | Kleist (1973, 194–216)                          |
| <i>La vie sainte Dieudonnée</i>                  | <i>Dieudonnée</i>         | Dirickx-Van der Straeten (1931, 159–91)         |
| <i>La vie saint Gregoire</i>                     | <i>Grégoire</i>           | Fant (1887)                                     |
| <i>Li livres de saint Jehan Ewangeliste</i>      | <i>Jehan Ewangeliste</i>  | Westberg (1943)                                 |
| <i>La vie saint Jehan Paulus</i>                 | <i>Jehan Paulus</i>       | Kleist (1973, 227–36)                           |
| <i>La vie du glorieux confesseur</i>             | <i>Confesseur</i>         | Kleist (1973, 294–302)                          |
| <i>La vie de saint Leu</i>                       | <i>Leu</i>                | Kleist (1973, 163–71)                           |
| <i>Les Poines d'Enfer</i>                        | <i>Vision</i>             | Kastner (1906, 427–50)                          |
| <i>La vie saint Sebastien</i>                    | <i>Sebastien</i>          | Micha (1971)                                    |
| <i>Un miracle de saint Servais</i>               | <i>Servais</i>            | Kleist (1973, 237–60)                           |
| <i>La vie de saint Thibaut confesseur</i>        | <i>Thibaut</i>            | Eastman Manning (1929)                          |
| <i>Un miracle de Thomas de Canterbury</i>        | <i>Thomas</i>             | Walberg (1922, 212–14)                          |
| <i>La vie saint Yves</i>                         | <i>Yves</i>               | Kleist (1973, 290–91)                           |
| <i>Le dit de Guillaume d'Angleterre</i>          | <i>Guillaume</i>          | Michel (1836–1840, 3: 173–211), Buzzetti (1990) |
| <i>Le dit de Robert le Diable</i>                | <i>Robert</i>             | Breul (1895, 464–509)                           |
| <i>Richard sans Peur</i>                         | <i>Richard</i>            | Conlon (1977)                                   |
| <i>Elegy de Troyes</i>                           | <i>Troyes</i>             | Fudeman (2008)                                  |
| <i>Le conte des vieillards tués</i>              | <i>Vieillards</i>         | Mussafia (1870)                                 |
| <i>La vie et l'istoire du mauvais riche home</i> | <i>Mauvais riche home</i> | Kleist (1973, 308–11)                           |
| <i>Du jeu des déz</i>                            | <i>Déz</i>                | Jubinal (1839–1842, 2: 229–34)                  |
| <i>Du roy qui avoit une amie</i>                 | <i>Roy amie</i>           | Jubinal (1839–1842, 2: 309–15)                  |
| <i>Du Dieu d'Amours, d'Esté et de May</i>        | <i>Dieu d'Amours</i>      | Raynaud (1885)                                  |
| <i>Des quatre sereurs</i>                        | <i>Quatre sereurs</i>     | Långfors (1932, 128–37)                         |

# INTRODUCTION

Que je croi qu'il n'est clers qui le peüst escrire  
Jehan de Saint-Quentin (?)

**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. Indeed, in the wake of the theoretical upheavals of the late twentieth century, most notably Roland Barthes's assertion that the author is dead (1968), and Michel Foucault's redefinition of the author as a function of discourse (1969, reprinted 1994), we may question why it matters at all to say that a text was written by a particular individual, especially when we know next to nothing about them.

Further, authorship in the Middle Ages was inevitably, in one way or another, a collaborative endeavour, with texts manipulated and changed by individuals who often operated centuries apart, as texts were copied and reworked. As Barbieri (2024, 17) notes, a manuscript text is rarely, if ever, just the work of a single author, but it is instead the cumulative output of "la pluralité des actants intervenant dans la confection (auteur, mais aussi copiste, compilateur, enlumineur, commanditaire, etc.)."<sup>1</sup>

It would also be a mistake to assume that there was one, monolithic concept of authorship that held across the Middle Ages. Early medieval vernacular literature can be characterized by its "auctorialité faible" (Barbieri 2024, 14), its "weak authorship," in which assertions of authorship, if made at all, were often secondary to the authority of the material itself. In the German heroic epic, for example, we find that anonymity was not just a trend but a genre-specific requirement, in which "Sagenkenntnis ist nicht individuelles Wissen eines Einzelnen, sondern kollektives Wissensgut einer gentilen Gesellschaft,"<sup>2</sup> and "[d]ie Darbietung des Sagenwissens in der Dichtung gilt [...] nach Stoff und Form nicht als schöpferische Leistung eines Einzelnen, der dafür namentlich genannt wird" (Hellgardt 1998, 46).<sup>3</sup> As Foucault

---

**1** "[t]he plurality of actors involved in the creation (author, but also copyist, compiler, illuminator, patron, etc.); translations (my own) will hereafter be given for all secondary material not written in English.

**2** "knowledge of legends is not the individual knowledge of a single person, but rather the collective knowledge of a gentile society."

**3** "[t]he presentation of legendary knowledge in poetry is [...] not considered, in terms of content and form, to be the creative achievement of an individual who is named for it."

himself noted for medieval literary texts (1994, 799–800): “leur anonymat ne faisait pas difficulté, leur ancienneté, vraie ou supposée, leur était une garantie suffisante.”<sup>4</sup> The author of such texts, even if named, is thus not an *auctor* but a *factor*, responsible for the production of a text but in no ways its ultimate originator.

From the thirteenth century onwards, by contrast, medieval readers developed an interest in the individual responsible for a text. This is most evident in the shift that occurs in biblical exegesis, from a focus on God as the spiritual inspiration for Scripture to a focus on the human behind the text as an *auctor* (Minnis 1988, 5), but it can also be observed in vernacular literature. Authorship is attributed from some of the earliest lyric poems in Old French, yet the key question of origins remains “for whom is it written?” until the fourteenth century, when it comes to be accompanied by a new question, namely “why is it written?” (Cerquiglini-Toulet 2006, 158). This latter question, in turn, brings about what Greene (2006a, 5) summarizes as “emergence of the ‘sentement’ [inspiration] as an authority from within, justifying the act of writing poetry independently of any patron.” As such, by this stage, “[t]he writer [of lyric poetry], even when he or she is commissioned, is no longer a servant or a minstrel, but is instead an author” (Cerquiglini-Toulet 2006, 162).

A similar diachronic trend has been observed in the use of the “I-narrator” in French chronicles and the verse romances (Marnette 2006b, 117). Uses of *je* to express the subjectivity of an author-narrator are absent from early examples, even those which contain named authors, and instead it is the prose form alone which serves to express “an impersonal, single, immanent truth” (2006b, 119). In texts dating from the end of the thirteenth century and later, by contrast, the “I-narrator” is a common feature which contributes to a “rhetoric of truth” (2006b, 118), in which “the I-narrator-author is engaged in the global coherence of his work” (2006b, 123), with the author thus playing a much more visible role than in earlier centuries.

By the fourteenth century, then, the *factor* has become an *auctor*, but great swathes of literary texts from this period remain unattributed, lacking references to named authors both within the texts themselves and in manuscript paratext. In such cases of anonymity, we may wonder whether authorship only mattered in certain contexts, and to certain people. Perhaps consistent attribution of authorship was simply less important to medieval readers than it is in our modern age with its “exceptionnel appétit d’auteur”

---

4 “their anonymity was not a cause for concern; their antiquity, whether real or supposed, was sufficient to guarantee their trustworthiness.”

(Zimmerman 2001, 7),<sup>5</sup> due both to the inherent complexity of authorship in the period, and the ways in which society considered and managed ideas surrounding intellectual property.

This unevenness gives rise to a conundrum: if attribution mattered only intermittently to medieval readers, what does modern scholarship stand to gain from attempting to identify the author of an anonymous medieval text? Besides the obvious human curiosity of wanting to discover the “maker” behind a work of art, literary or otherwise, textual philology has long been concerned with questions of authorship. As Nichols (2006, 79) argues, if Heidegger (1975, 17) was right to argue that the artist is the origin of the work and the work is the origin of the artist, then “textual philology requires an author” for “without an author, there can be no philology.”

Seen in this light, the uneven status of medieval authorship does not invalidate attempts at attribution, but rather forces us to reconsider what is at stake when we attribute a medieval text to an author.<sup>6</sup> It is the thesis of this book that authorial attribution can serve as a valuable 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 corpus, revealing the similarities shared by a given group. At the same time, to assert that a corpus is the work of one author is also to make an exclusionary claim, differentiating 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 approach to reading 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 book, I show that digital and statistical approaches bespoke to the content and challenges of a particular corpus have much to offer the examination of uncertain authorship. I take a corpus of texts found mainly

---

**5** “exceptional appetite for authorship.”

**6** For a fuller account of medieval conceptions of authorship, and of the modern theories that have shaped their study, see Julien Stout (2025), esp. pp. 1–102.

within a single manuscript, namely Paris, Bibliothèque nationale de France, MS français 24432 (henceforth fr. 24432), which was compiled around the middle of the fourteenth century. I examine the case for the attribution of this corpus to one Jehan de Saint-Quentin, whose name appears in a single text of the corpus.

Through this analysis, I reinforce the attribution of twenty-four texts to Jehan made by previous scholarship, and I provide evidence for the novel conclusion that there are further works which deserve inclusion in his *œuvre*, specifically *Guillaume* and *Robert*,<sup>7</sup> both of which appear in fr. 24432. I also suggest that it is possible to identify two echoes of earlier texts written by Jehan, which are now lost, in the shape of *Richard* and *Jehan Paulus*. Collectively, I term these claims the “shared-author hypothesis.”

The main methodological contribution of this study, meanwhile, is the suggestion that methods of digital, statistical, and codicological analysis are at their most effective when used in a bespoke way alongside the linguistic and textual approaches common in previous scholarship. This allows for the combination of a humanistic understanding of the object of study, founded on close textual engagement, with concrete, reproducible results, which provides greater clarity than either approach in isolation. The work presented here thus 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 book, and reshaping the ways we approach medieval authorship.

## Jehan de Saint-Quentin and fr. 24432

The central manuscript of this study, fr. 24432, 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 ninety items in verse and prose, including a few additions by the manuscript’s later readers, as well as a table of contents.

The manuscript appears within well-studied groupings of codices, as manuscript *P* of the *fabliaux*,<sup>8</sup> manuscript *Q* of the works of Marie de France,<sup>9</sup> and manuscript *D* of the works of Rutebeuf.<sup>10</sup> It is, however, peripheral to the

---

**7** For short titles by which I will refer to these and other texts throughout, see the lists in the frontmatter.

**8** See, for example, Noomen and van den Boogaard (1983–1998).

**9** On which, see Busby (2011).

**10** Zink (1989–1990).

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, it is quite late in the transmission cycle of all three.

Amongst the manuscript's less well-known texts is a partially contiguous collection of short didactic narratives, written in alexandrine monorhyme quatrains (that is, four-line stanzas of twelve-syllable lines, all of which share a rhyme). In most cases, they tell stories of sinners who are driven to the most appalling acts, including adultery, incest, Host desecration, pillage, infanticide, and uxoricide, but who are given opportunities to change their ways and be saved. These poems have no stated author within either the text itself or the manuscript paratexts, with one exception: in line 221 of *Chevalier et escuier* (fols. 112ra–115rb), an author-narrator refers to himself as “Jehan de Saint-Quentin.” This is the only time this name appears in fr. 24432, 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).<sup>11</sup> 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. However, as Kedar (1973) has noted, medieval toponymic bynames are notoriously unreliable as a method of determining a person's place of origin, and there is no evidence that Jehan spent any time at all in Saint-Quentin. 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<sup>e</sup> siècle,”<sup>12</sup> given that our best guess at dating *Chevalier et escuier* comes only from the simple fact that most of the datable texts within fr. 24432 fall within that period.<sup>13</sup>

Jehan de Saint-Quentin would therefore appear to be an almost wholly unknowable figure, an author lost 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

---

**11** 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.

**12** “first half of the 14<sup>th</sup> century.”

**13** For further discussion of questions of localization and dating, see below.

the last century which suggests that Jehan's contribution is more significant, and that he ought to be recognized as the author of other poems within fr. 24432, as well as of some in other manuscripts.

The principal aim of this study is therefore 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 and others. In each chapter, I examine a given 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 bespoke digital and statistical techniques, developed to fit a particular corpus of texts, as a means of answering diverse research questions regarding the attribution of authorship.

## Existing Attribution to Jehan de Saint-Quentin

Before we begin 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 *Chevalier et 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 twenty poems, all found within fr. 24432, which are listed in Table 1. Gröber was primarily interested in the texts' sources, and gave no evidence or explanation for his inclusion and non-inclusion of any texts.

Following this, Joseph Morawski published an analysis of a corpus of twenty-two poems which, he claimed, could all be attributed to Jehan de Saint-Quentin, whom he termed a “jongleur oublié” (1939, 327).<sup>14</sup> Morawski's list differed from that of Gröber through the inclusion of an extra poem from within fr. 24432, *Merlin*, as well as two from outside the manuscript (*Abesse* and *Sauveur*). He also excluded *Florence*, which Gröber had considered worthy of inclusion in his corpus. The list of poems which Morawski attributes to Jehan can be found in Table 2.

---

14 “forgotten *jongleur*.”

Table 1. Texts attributed to Jehan de Saint-Quentin by Gröber (1888–1902).

| Short Title                 | Folios in fr. 24432 |
|-----------------------------|---------------------|
| <i>Respon</i>               | 14rb–17ra           |
| <i>Chanoines</i>            | 36vb–42ra           |
| <i>Pommes</i>               | 53va–56vb           |
| <i>Romme</i>                | 99rb–102rb          |
| <i>Deux chevaliers</i>      | 104va–107rb         |
| <i>Enfant rosti</i>         | 107rb–109vb         |
| <i>Povre chevalier</i>      | 109vb–112ra         |
| <i>Chevalier et escuier</i> | 112ra–115rb         |
| <i>Narbonne</i>             | 115rb–118rb         |
| <i>Chevalier hermite</i>    | 118rb–120va         |
| <i>Cordouanier</i>          | 120va–122vb         |
| <i>Juitel</i>               | 122vb–125ra         |
| <i>Enfant sauva mere</i>    | 125ra–128ra         |
| <i>Eaue et vergier</i>      | 128ra–130rb         |
| <i>Pain</i>                 | 130rb–132va         |
| <i>Chien et mescreant</i>   | 132va–135rb         |
| <i>Pecherresse</i>          | 135rb–137rb         |
| <i>Florence</i>             | 215rb–126va         |
| <i>Buef</i>                 | 247ra–257va         |
| <i>Beguine</i>              | 312vb–315va         |

Table 2. Texts attributed to Jehan de Saint-Quentin by Morawski (1939, 328–30).

| Short Title                 | Folios in fr. 24432       | Morawski's ref. |
|-----------------------------|---------------------------|-----------------|
| <i>Respon</i>               | 14rb–17ra                 | V               |
| <i>Chanoines</i>            | 36vb–42ra                 | XX              |
| <i>Pommes</i>               | 53va–56vb                 | XIX             |
| <i>Romme</i>                | 99rb–102rb                | VII             |
| <i>Deux chevaliers</i>      | 104va–107rb               | II              |
| <i>Enfant rosti</i>         | 107rb–109vb               | XIII            |
| <i>Povre chevalier</i>      | 109vb–112ra               | III             |
| <i>Chevalier et escuier</i> | 112ra–115rb               | I               |
| <i>Narbonne</i>             | 115rb–118rb               | VIII            |
| <i>Chevalier hermite</i>    | 118rb–120va               | IV              |
| <i>Cordouanier</i>          | 120va–122vb               | VI              |
| <i>Juitel</i>               | 122vb–125ra               | XIV             |
| <i>Enfant sauva mere</i>    | 125ra–128ra               | XV              |
| <i>Eaue et vergier</i>      | 128ra–130rb               | XVI             |
| <i>Pain</i>                 | 130rb–132va               | XVII            |
| <i>Chien et mescreant</i>   | 132va–135rb               | XVIII           |
| <i>Pecherresse</i>          | 135rb–137rb               | X               |
| <i>Merlin</i>               | 199ra–201vb               | XII             |
| <i>Buef</i>                 | 247ra–257va               | XXI             |
| <i>Beguine</i>              | 312vb–315va               | XI              |
| <i>Sauveur</i>              | Not present <sup>15</sup> | XXII            |
| <i>Abesse</i>               | Not present <sup>16</sup> | IX              |

**15** Appears only in Paris, Bibliothèque nationale de France, MS Arsenal 2115, fols. 48r–48v, 52r–57v.

**16** Appears only in Paris, Bibliothèque nationale de France, MS français 12483, fols. 66ra–68rb.

Unlike Gröber, Morawski was explicit in his exclusion of texts which, he argued, shared notable similarities with the poems in his corpus, but were not sufficiently similar to be considered the works of Jehan. In his 1939 work, he explained his exclusion of *Florence* and *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, 351–52).<sup>17</sup> 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.<sup>18</sup> In 1940, Morawski returned to the topic of the corpus of Jehan's poems, but did not add any further texts, instead focusing on the explicit exclusion of another four texts from the corpus,<sup>19</sup> and reinforcing his conclusion that *Anelés* and *Florence* are not the works of Jehan. Morawski reaches his conclusions through a largely linguistic approach examining the use of possessive adjectives, to which we will return in detail below.

Birger Munk Olsen (1963a, 1963b, 1978) was the next to take up and examine Morawski's claims regarding Jehan's authorship of these poems, and the most recent to do so in detail. He is largely critical of Morawski's approach to attribution, but does not take issue with most of his decisions to exclude certain texts. Munk Olsen's work culminates in his 1978 monograph, the first attempt at a complete edition of Jehan's *œuvre*, which contains a total of twenty-four poems, including all of those cited by Morawski, with the addition of *Anelés* and *Florence*, the two poems which Morawski had believed were 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

---

**17** “we believe that these two *Dit-romances* are the work of the same author, but we have not reached the conviction that they are by Jean.”

**18** Specifically, the alexandrine quatrain versions of the lives of Saint Alexis, Saint Dieudonnée, and Saint Gregory, included as *Alexis*, *Dieudonnée*, and *Grégoire* in the corpus of comparable texts which we will introduce below as the “A-corpus.” For editions, see Paris and Pannier (1872), Dirickx-Van der Straeten (1931), and Fant (1887) respectively.

**19** *Basile*, *Servais*, *Jehan Paulus*, and *Richard*. For an edition of the first text, see Denomy (1956), for the second and third, see Kleist (1973), and for the fourth, see Conlon (1977). We will argue later in our analysis that *Jehan Paulus* and *Richard* do, in fact, deserve partial inclusion in Jehan's *œuvre* as echoes of earlier texts which were written by Jehan and later reworked. Morawski also briefly mentions *Robert*, a text which appears within fr. 24432, but makes no claim as to whether it is the work of Jehan. We will also return to this text later in our analysis, arguing that it merits straightforward attribution.

we return to throughout this study, 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 various factors, he also rejects the possibility that *Alexis* could be the work of Jehan. Munk Olsen's 1978 list can be found in Table 3.

Since this work, very little attention has been paid to the attribution of these poems, with references to Jehan's *œuvre* generally taking Munk Olsen's corpus as the definitive list.<sup>20</sup> 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, 96n6) questions why *Robert* 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 fr. 24432, and "ressemble beaucoup aux autres dits."<sup>21</sup> He makes no mention of the attribution of *Guillaume*, the other unattributed text within fr. 24432 which fits the same criteria, and argues that the similarities across all of Jehan's texts in "la versification, le style, la fonction du récit [...] relève[nt] davantage du genre

Table 3. Texts attributed to Jehan de Saint-Quentin by Munk Olsen (1978).

| Title                       | Folios in fr. 24432 | Munk Olsen's ref. |
|-----------------------------|---------------------|-------------------|
| <i>Respon</i>               | 14rb–17ra           | A                 |
| <i>Chanoïnes</i>            | 36vb–42ra           | B                 |
| <i>Pommes</i>               | 53va–56vb           | C                 |
| <i>Romme</i>                | 99rb–102rb          | D                 |
| <i>Deux chevaliers</i>      | 104va–107rb         | E                 |
| <i>Enfant rosti</i>         | 107rb–109vb         | F                 |
| <i>Povre chevalier</i>      | 109vb–112ra         | G                 |
| <i>Chevalier et escuier</i> | 112ra–115rb         | H                 |
| <i>Narbonne</i>             | 115rb–118rb         | I                 |
| <i>Chevalier hermite</i>    | 118rb–120va         | J                 |
| <i>Cordouanier</i>          | 120va–122vb         | K                 |
| <i>Juitel</i>               | 122vb–125ra         | L                 |
| <i>Enfant sauva mere</i>    | 125ra–128ra         | M                 |
| <i>Eaue et vergier</i>      | 128ra–130rb         | N                 |
| <i>Pain</i>                 | 130rb–132va         | O                 |
| <i>Mescreant</i>            | 132va–135rb         | P                 |
| <i>Pecherresse</i>          | 135rb–137rb         | Q                 |
| <i>Merlin</i>               | 199ra–201vb         | R                 |
| <i>Florence</i>             | 215rb–226va         | S                 |
| <i>Anelés</i>               | 231ra–241rb         | T                 |
| <i>Buef</i>                 | 247ra–257va         | U                 |
| <i>Beguine</i>              | 312vb–315va         | V                 |
| <i>Sauveur</i>              | Not present         | W                 |
| <i>Abesse</i>               | Not present         | X                 |

<sup>20</sup> See, for example, Simpson (2019) and Cerquiglini-Toulet (2022).

<sup>21</sup> "is very similar to the other *dits*."

littéraire que de l'auctoralité" (2022, 96),<sup>22</sup> a claim which we will dispute on the basis of comparisons to other texts of a similar type. Cholette is nevertheless 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."<sup>23</sup> While collective compilation is a mark in favour of the shared-author hypothesis, we must go much further in order to be able to say with any degree of certainty that these poems can be described as the work of a single author, and to determine which others might also be included.

## Localization and Dating

Much of the possible linguistic analysis of the Jehan's *œuvre* 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 localizing 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*."<sup>24</sup> At best, it seems, Jehan's language can be best described as containing a mixture of traits associated with several different, predominantly northern and western, dialects.

Morawski (1939, 333) categorizes the language of the *J-corpus* as being primarily a "mélange de formes franciennes avec des formes septentrionales et occidentales,"<sup>25</sup> which he notes is a common feature of many of the French poems from this period written in alexandrine rhyming quatrains. Morawski's evidence in support of this claim 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.

---

**22** "the versification, style, and function of the narrative [...] pertain more to the literary genre than to authorship"

**23** "the fact that the narrative *dits* were grouped together in the same manuscript is not sufficient to affirm that they were written by the same author."

**24** "based on linguistic features attested as authentic by rhyme or scansion, it is quite difficult, if not impossible, to precisely localize the author of the *Dits*."

**25** "mixture of Francien forms with northern and western forms."

106) and the interchangeable use of the 2.PL possessives *vo* and *vostre* (1939, 333).<sup>26</sup> Munk Olsen observes further Picard traits in Jehan's work, 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 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*.<sup>27</sup>

That said, as Munk Olsen (1963a, 124) observes, these picardisms in Jehan's *œuvre* are the exception to the rule, since "la plupart des particularités du texte sont caractéristiques de l'Ouest."<sup>28</sup> 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. 670–71), monophthongization of pre-consonantal /ai/ to /ɛ/ (as in *nestre* rhyming with *prestre*; *Narbonne*, vv. 144–45), and the raising of tonic /o/ to /u/ (as in the rhyme of *seignor* with *dolour*; *Chevalier et escuier*, vv. 217–18), alongside the appearance of forms such as *el*, *li*, and *celi*.<sup>29</sup> It is of course possible that many of these forms could result from scribal changes, but their preponderance in circumstances where such a change would have resulted in inconsistencies in the rhyme or metre allows us to say with some certainty that this is not always the case, and that the language of Jehan's texts thus exists somewhere between Picard, western, and *francien* varieties.

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.

The best solution to this inextricable problem would certainly be to see it as an artificial and hybrid language: the author took notes in any dialect [...],

**26** 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.

**27** We also observe such Picardisms within the *A-corpus*, introduced below. See, for example, *parlon* (*Richard*, v. 321), *perdera* (*Quatre sereurs*, v. 105), and *fiex* (*Alexis*, e.g., v. 41).

**28** "most of the peculiarities of the text are characteristic of the west [of France]."

**29** Cf., in our *A-corpus*, trisyllabic *eslongnierent* rhyming with *entrerent* (*Guillaume*, v. 386–87), *neistre* rhyming with *pestre* (*Jehan Ewangeliste*, v. 226), and *doctor* rhyming with *dolour* (*Vision*, v. 37–38).

and since he had no sense of a normalized language, he had no qualms about using them even if they were in contradiction with his own dialect, which could have been Picard just as well as one of those from the west-central part of France.

There are a great many questionable assumptions in this claim, not least that Jehan would have had no concept of a normalized, 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 attributed to Jehan, 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 the purposes of this study, namely that the texts of Jehan's *œuvre* present a degree of unity in the extent to which they combine Picard, western, and *francien* forms, regardless of the reasons why such a combination exists. This does not provide conclusive evidence for the shared-author hypothesis, given the extent to which such forms are also found elsewhere, but at the very least it does not discourage collective attribution.

As for the dates of the texts attributed to Jehan, there is little concrete evidence to come from the language itself. 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 fr. 24432, although these are not directly linked any of the texts attributed to Jehan.<sup>30</sup> To support this linguistically, he notes that analogical forms such as non-epicene *grande* occasionally appear in rhyming position or at a moment where the two-syllable form is metrically necessary (once each in *Narbonne*, *Chevalier et escuier*, and *Buef*), 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 guère plus avancés quoique nous n'ayons que le choix entre la fin du XIII<sup>e</sup> et le début du XIV<sup>e</sup>

---

**30** The following dates are given within the manuscript's texts: 1332 (fol. 245ra), 1338 (fol. 17rb), and April 23, 1345 (fol. 266rb). The following dates are given in the surrounding paratext: 1313 (fol. 146va), 1318 (fol. 148rb), 1320 (fol. 152ra), and 1324 (fol. 142vb). The *terminus post quem* for the manuscript's production is therefore St. George's Day 1345, April 23. However, due to the range of dates given in paratexts (the earliest being 1318), as well as the appearance of much earlier texts such as Marie de France's *Yonec*, likely dating to the late twelfth century (Busby 2011), it seems that a narrow dating based on the manuscript's other texts is impossible.

siècle" (1963a, 125).<sup>31</sup> Evidence of forms associated with the late thirteenth century comes in the form of enclisis in *nel*, *nes*, and *sel*,<sup>32</sup> 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, where this can be confidently determined to be the result of an authorial choice, rather than a scribal one (that is, where nominative and oblique forms are imparisyllabic).

Ultimately, then, while there are trends to be observed within Jehan’s language which increase the unity of his corpus, these do not offer much clarity as to where or when he was writing, nor do they differ especially from other, comparable texts. A more comprehensive, holistic approach to the corpus is therefore required to assess the validity of the shared-author hypothesis.

## Four New Additions to Jehan’s *œuvre*

As discussed above, the main contribution of this book to the study of Jehan de Saint-Quentin’s work is what I have termed the “shared author hypothesis.” That is, my central claim is that, alongside the continued attribution of the twenty-four texts previously attributed to Jehan, it is possible to identify two further texts which merit inclusion in his *œuvre*, namely *Guillaume* and *Robert*, and two which survive as reworkings of earlier texts by the same author, namely *Richard* and *Jehan Paulus*.

Given the agglutinative nature of the arguments presented throughout this study, it is not possible here to set out the full case for their addition to the *œuvre*. At this stage, I seek only to introduce these texts, and to examine previous scholarly attention that they have received, to highlight the significance that they will, by the end of the study, come to hold as works (or echoes of works) by Jehan.<sup>33</sup>

**31** “we are hardly any further along, although we only have the choice between the end of the thirteenth and the beginning of the fourteenth century.”

**32** Also observed in *Robert* (vv. 568, 675).

**33** All four texts are included in our *A-corpus* of narrative poems written in alexandrine monorhyme quatrains that have hitherto not been attributed to Jehan,

*Robert* is mentioned by Morawski (1940, 518), but only as a comparator for *Richard*, with no attention paid to whether the text should be attributed to Jehan. As we have already noted, Cholette (2022, 96n6) observes that *Robert* “ressemble beaucoup aux autres dits,”<sup>34</sup> and suggests that the text could be the work of Jehan, but offers no additional evidence to support this claim. The text appears in fr. 24432, immediately after *Merlin* and immediately before *Florence*, both texts previously attributed to Jehan, and thus appears a prime candidate for consideration based on its compilation alone.<sup>35</sup> It also appears in a further three manuscripts, namely Besançon, Bibliothèque municipale, MS 588; Paris, Bibliothèque nationale de France, MS français 1881; and Paris, Bibliothèque nationale de France, MS français 12604. The poem aligns closely with the didactic content of the *J-corpus*, telling 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.

Turning to *Guillaume*, this text also shares a great many similarities with Jehan’s texts, both in form and content. Alongside its appearance as the first item of fr. 24432, *Guillaume* also appears in an early fourteenth-century manuscript London, British Library, MS Additional 15606, making the copy in fr. 24432 likely not the earliest surviving witness of the text. The poem is again simultaneously narrative and 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.

While *Guillaume* and *Robert*, as we shall see, can be straightforwardly considered the work of Jehan, the same is not true for *Richard*. Instead, the text that survives to the present day is most likely a reworking of an original version composed by Jehan. It is preserved in a sixteenth-century manuscript Paris, Bibliothèque nationale de France, MS nouvelle acquisition française 10721, and in a late fifteenth-century incunable edition with the incipit: *S’ensuyt le romant de Richart filz de Robert le diable qui fut duc de Normendie*. The poem tells the story of Richard, the son of the eponymous (anti-)hero of *Robert*, who proves through various trials that he deserves

---

as discussed below. Initially, at least, they are thus compared with texts previously attributed to our author, and they will be singled out for specific comment as the evidence in favour of the shared-author hypothesis begins to grow.

**34** “is very similar to the other *dits*.”

**35** The text appears at fols. 202ra–215rb between *Merlin* (fols. 199ra–201vb) and *Florence* (fols. 215rb–226va).

the title of *sans peur*. Morawski (1940, 516) suggests that at least a portion of *Richard* dates to the fourteenth century, and he observes points of intersection between *Robert* and *Richard* which take on additional significance in the context of our shared-author hypothesis: in the closing lines of both poems, the narrator(s) provide information about the character Richard, in the former case proceeding down the family tree from his father Robert:

Le fille l'emperere ot de li un enfant  
C'on appella en France dant Richart le Normant,  
Qui fist moult de prouesce tant comme il fu vivant;  
De Fesquan l'abaie fonda, je vous creant.

L'abbaye et le moustier de Fescamp il fonda  
Et Saint Wandrille aussi. Sainte vie mena.  
Avecques Charlemagne oultre les mons passa  
Et fist de grans proesses en tant comme il dura.

Avecques Karlemaigne passa outre la mer.

(*Richard*, vv. 807–10)

(*Robert*, vv. 1009–13)<sup>36</sup>

The comparable references to Charlemagne and Fécamp abbey, and the use of parallel structures such as “passa outre la mer” with “oultre les mons passa,” and “Qui fist moult de prouesce tant comme il fu vivant” with “Et fist de grans proesses en tant comme il dura” are all suggestive of significant influence, if not outright copying, in earlier versions of *Richard*, providing initial evidence in support of the suggestion that these two poems could, in their putative earliest versions, be considered the work of one author. In this context, it is worth considering the opinion of Conlon (1977, 33), the most recent editor of the text, who claims of the poem in its current form:

[*Richard*] was written in Paris and dates from the closing years of the fifteenth century or the first half of the sixteenth. This does not exclude the existence [sic] of an earlier version written in a dialect from Upper Normandy, but there is no conclusive evidence for such a version in the extant text.

Contrary to Conlon's suggestion, the evidence which we shall accumulate over the course of this study, points strongly towards the possibility that an earlier version of *Richard* did exist, and that it was written by Jehan.

**36** Note that when giving parenthetical citations for texts of the *J-* or *A-corpora*, the two main corpora used in this study which are defined in Chapter 1, I am not directly citing any specific edition. Instead, the citation is of my own textual data which I have produced with reference to one or more existing editions, as well as my own reading of the manuscript sources. Here, quotation marks are used within citations of source material in text and in tables to indicate direct discourse, i.e., as examples of “quotes within quotes.” For details of the development of the data for the *J-corpus*, see the discussion that follows, and for the editions consulted for the development of the *A-corpus*, see the list of short titles in the frontmatter.

As with *Richard*, *Jehan Paulus* shares sufficient similarities with Jehan's other poems to suggest that it deserves inclusion in his *œuvre*, but the version which survives today is likely a reworking of an earlier version by Jehan. The poem is found only in Grenoble, Bibliothèque municipale, MS 319, the fifteenth-century manuscript which also contains the only other surviving copy of *Florence*. The appearance of this text near *Florence* is already an indicator that the texts are closely related, perhaps indicating that *Jehan Paulus* was reworked on the road to inclusion in the manuscript, while *Florence* escaped the same fate, at least to such a significant extent. The poem recounts a tale which, while unusual in that it is explicitly a saint's *vita* (unlike all but *Sauveur* within the texts previously attributed to Jehan), fits almost exactly with Jehan's apparent narrative fixations, particularly in his shorter texts: Devil-induced debauchery and violence, penance, and a miracle performed by Mary.<sup>37</sup> The eponymous saint, himself a hermit, is tricked into lying with and murdering a young noblewoman, who is kept alive only by the grace of the Blessed Virgin and a prayer that the hermit repeats incessantly.

None of these four texts are perfect archetypes of Jehan's *œuvre*. As this study will set out to prove, however, they do share sufficient similarities, both as a group and with works previously attributed to Jehan, that the shared-author hypothesis is much more plausible than previous scholarship has suggested.

---

**37** Indeed Morawski (194, 513–16), while still arguing that the text is not the work of Jehan, notes that the Marian focus in *Jehan Paulus* is augmented compared to other retellings of the same story.

## METHODOLOGY

PROGRESS IN RECENT decades on digital and statistical approaches to the analysis of text allows us to approach Jehan's *œuvre* from directions unavailable to previous scholarship. So many of Morawski and Munk Olsen's arguments are founded on classically philological, qualitative claims. Such approaches certainly have a value, yet as we shall see throughout our analysis, they often lack the reliability and reproducibility that more modern approaches can offer. The aim of this study is therefore not to restart examination of Jehan's poems from the ground up, rather it is to build on existing work through the use of digital and statistical approaches, 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 complement earlier arguments, building on scholars' work using methods that were unavailable to them to critically analyse the conclusions they reach. In others, particularly in cases where scholarly thought has moved on in the half-century since the most recent large-scale work on Jehan's poetry, I offer alternative approaches to the examination of the corpus to those which have been attempted previously.

Many of the methods in this study overlap with the field of stylometry, as pioneered by Frederick Mosteller and David L. Wallace (1963). For example, the examination of line-initial words on pp. 118–28 shares similarities with so-called “unmasking” approaches (Argamon 2018, 14). In the decades since Mosteller and Wallace's initial experiments, stylometry has benefitted from significant advances in computational capacity and statistical research, and has produced promising results within both literary studies and beyond,<sup>1</sup> and yet the purpose of this monograph is not simply to apply “off-the-shelf” stylometric techniques to Jehan's poems in the hope of arriving at an immediate and definitive answer to the question of attribution. Rather, the techniques and approaches presented here are predicated on the contention

---

<sup>1</sup> On recent developments in the literary uses of stylometry see, for example, Craig (2021, 2023), and on the role of stylometry as a tool of forensic analysis, see Argamon (2018). Eder et al. (2016) presents an R package containing tools that can be applied to questions of digital stylometry.

that the study of medieval authorship, fragile and elusive as it is, requires a more nuanced methodological framework than might be applicable in other contexts. Where many stylometric approaches treat a given text first and foremost as a mere “bag of words,” within this study I adopt explicitly bespoke, text-focused, and data-driven methods, which I have developed to suit the linguistic, formal, and codicological specificities of Jehan’s poems, and which are grounded in an in-depth, humanistic understanding of the poems’ content and historical context.

This study is thus the first to approach Jehan’s authorship using digital techniques, and the first to do so in a bespoke way that offers a model for similar cases of uncertain authorship. Each chapter approaches the same question from different angles, using carefully selected techniques to shed further light on the shared-author hypothesis. As might be expected, no one method alone can be seen to provide a “smoking gun” which proves or disproves Jehan’s authorship of a given text with certainty. Instead, the chapters proceed in an agglutinative way, gradually accreting evidence which can then be taken holistically to reach our overall conclusions.

Given the cumulative nature of this approach, at the conclusion of each main chapter I present two distinct sections which set out that chapter’s contributions, as a means of consolidating and assessing the accumulated evidence as we progress through the study. In the first section, “Building the Case for the Shared-Author Hypothesis,” I summarize what the chapter has contributed to our agglutinative understanding of Jehan’s *œuvre*, both in terms of the coherence of the *J-corpus* of texts traditionally attributed to Jehan, and the validity of the attribution of our four additional texts to the same author. The “Methodological Contribution” section, meanwhile, abstracts away from the specific context of Jehan de Saint-Quentin, and considers the utility of the bespoke methods and approaches presented in the chapter to broader questions of uncertain authorship.

## Use of Digital and Statistical Methods

For the digital and statistical methods applied in this study to be accessible to philologists and digital humanists alike, it is also necessary first to set out and explain the technological and mathematical frameworks that will underpin our analysis.

The core analysis of the target corpus of Jehan’s texts, named below as the *J-corpus*, proceeds for the most part from a single digital encoding of fr. 24432. The chosen format of this encoding is TEI-XML, the Text Encoding Initiative (TEI Consortium 2025), a variant of the encoding language XML

specifically designed to allow text of all kinds, including medieval texts, to be structured in a machine-readable way. The resulting encoding of fr. 24432 is therefore not merely a single transcription but a complex data structure, which contains variants of different textual elements including both expanded and unexpanded abbreviations, regularized and unregularized word forms, and corrected and uncorrected errors.<sup>2</sup>

The benefit of such an approach is that decisions to expand, regularize, and correct text do not therefore have to be made during the process of encoding, and can instead be determined on a case-by-case basis when data are later extracted from it. In this case, I do so using a combination of scripts written in XSLT (Extensible Stylesheet Language Transformations; W3C 1999) and R (R Core Team 2025) to perform transformations that extract the plain text files suitable for computational analysis from the core TEI-XML file.

When searching for specific patterns within these processed text files, I employ the formal expression language known as regular expressions (regex) to isolate specific strings of characters. This proves particularly important when examining features such as the use of masculine and feminine rhymes discussed on pp. 89–95, and so-called “easy” rhymes discussed on pp. 96–102, where specific orthographic patterns must be identified amidst the inherent variability of medieval French spelling. Without seeking to provide a full account of regex notation,<sup>3</sup> suffice it to say here that the uses of regular expressions in this study generally centre around the identification of variant forms of the same underlying linguistic element, such as word-final *-ant* and *-ans*, which could be represented in regex notation as `an[ts]$,` where square brackets indicate a choice between two or more characters, and the dollar sign indicates the end of a string of characters.

To measure the relationships between different textual variables, I rely on three core statistical tests. The *Pearson’s product-moment correlation test*, as established in Pearson (1895) and discussed Cohen et al. (2013, §2.2), is used to calculate the strength of a relationship between two continuous variables. For instance, on p. 35, I use this test to determine the relationship between text length and prologue length within the texts previously attributed to Jehan. Following conventional practice, when statistical correlation is calculated for a given feature, it is described as strong when the correlation coefficient is between  $\pm 0.5$  and 1, and weak when it is between  $\pm 0.25$

**2** For access to the full encoding, alongside XML, TXT, CSV, and HTML outputs with an accompanying IIIF viewer, see Dows-Miller (2025b).

**3** For such a resource, see Friedl (2006).

and 0.5. Positive correlation coefficients indicate a positive correlation (that is, as one variable increases, so does the other), while negative correlation coefficients point to a negative correlation (where, as one variable increases, the other decreases, and *vice versa*).

Second, when comparing the average rates of a feature across two different groups, such as the rate of so-called “easy” rhymes in Jehan’s texts compared to a reference corpus (as discussed on pp. 101–2), the *Welch Two Sample t-test*, as established in Welch (1947) and discussed in Ruxton (2006), is used to determine whether the difference between the two means (averages) is significant. This therefore enables us to determine whether the difference between the means of both corpora exceeds what would be expected by chance, or whether they are instead more comparable.

Finally, in cases where we consider the use of a given feature in a single text and compare it to that observed in the remainder of a corpus, for example the rates of appearance of individual words at the beginning of a line as discussed on pp. 122–23, the observed and expected values are often too small to allow for the use of a Chi-squared “goodness of fit” test, as would otherwise be standard. In such cases, some (such as Campbell 2007) have recommended a combined or conditional approach drawing on both the Chi-squared and Fisher-Irwin tests, but I have ultimately chosen where possible to make use of a bespoke randomization test, due to both its mathematical simplicity and the methodological uniformity which it allows. The full R script necessary to generate the function is given in Figure 1.

In short, this randomization method compares the observed appearances of a feature between a target corpus and a reference corpus, to determine whether any observed difference is statistically significant. It does so by taking the number of observed instances and total counts for both the target and reference corpora, and then calculating a *null rate*, which represents the expected probability of observing the feature were there to be no difference between the two corpora in the rate at which this feature appears. It then generates a million samples for both corpora using a binomial distribution, which allows us to calculate how likely it would be to observe a difference at least as extreme as the one observed, were both corpora to share the same underlying null rate. The randomized nature of this test means that this final probability will always fluctuate somewhat, but not enough to vary the result of the test.

The benchmark for statistical significance in all three tests, by which we can affirm the validity of our claims that compared rates across and within corpora are notably distinct, is the *p-value*. This represents the probability that the observed results occurred under the *null hypothesis*, that is, that the

```

1  #randomised p_value tests for target vs reference corpus
2  p_value_rand <- function(
3    observed_tar = integer(),
4    total_tar = integer(),
5    observed_ref = integer(),
6    total_ref = integer(),
7    sample_size = 1000000) {
8
9    sum_observed <- observed_tar + observed_ref #sum of observed instances
10   sum_total <- total_tar + total_ref #sum of total instances
11
12   #null rate: expected rate if there is no difference between a and b
13   null_rate <- sum_observed / sum_total
14
15   #observed difference between probabilities
16   observed_prob_dif <- (observed_ref / total_ref) - (observed_tar / total_tar)
17
18   #generate random samples, counting observed instances
19   ref_rand <- rbinom(sample_size, total_ref, null_rate)
20   tar_rand <- rbinom(sample_size, total_tar, null_rate)
21
22   #work out differences in probability between two samples
23   rand_prob_difs <- (ref_rand / total_ref) - (tar_rand / total_tar)
24
25   #make differences in probability absolute
26   abs_rand_prob_difs <- abs(rand_prob_difs)
27   abs_observed_prob_dif <- abs(observed_prob_dif)
28
29   #calculate the p value (the proportion of the random sample for which
30     #the difference in probability exceeds the observed probability)
31   p_value <- mean(abs_rand_prob_difs >= abs_observed_prob_dif)
32
33   p_value
34 }

```

Figure 1. Custom R function to calculate a p-value for the use of a given feature between a target and reference corpus using randomized sampling.

rate of occurrence of an observed feature of the target corpus does not differ from that of the same feature in the reference corpus. Where necessary, p-values are given to five decimal places, and following standard practice I set the threshold for statistical significance at  $p < 0.05$ . That is, for an observation to be statistically significant, the probability of observing a result at least as extreme as the one obtained, were there to be no underlying difference between the two corpora, must be less than 5%. A p-value lower than this threshold allows us to reject the null hypothesis and suggest with confidence that an observed feature, for example the use of *tous et toutes* discussed on pp. 71–73, is indeed a distinguishing characteristic of Jehan's style. By accreting these statistically significant results, we move away from subjective impressions of authorial preferences and styles towards a holistic, data-driven examination of the shared-author hypothesis.

## Target Corpus

As the central object of study for this study, I take the texts contained within Munk Olsen's (1978) corpus of Jehan's poems, a corpus of some 63,347 words which I term the *J-corpus*. This corpus includes both the twenty-two texts attributed to Jehan that appear within fr. 24432, as well as the two (*Sauveur* and *Abesse*) found outside it. While we will critically examine the boundaries of this corpus, it serves 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.

When compiling the texts of the *J-corpus*, I have paid particular attention to Munk Olsen's (1978) edition of the poems he attributes to Jehan, while also referring directly to the manuscript to ensure that prior editorial preferences do not cloud our conclusions. In most cases for the analysis in this study, plain text data akin to a single-manuscript critical edition have sufficed. As described above, this has been extracted from a TEI-XML encoding of fr. 24432 through an XSLT transformation, which expands abbreviation and accepts appropriate additions, deletions, regularizations, 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 also appears outside of fr. 24432, and it does not contain modernized 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 fr. 24432, *Sauveur* and *Abesse*, Munk Olsen's (1978) edition has sufficed.

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 examine in detail in Chapter 5. 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 (in most cases both are true), they have been lemmatized. That is, we assign an arbitrarily chosen "dictionary" form to each instance of a given lexeme, to extract clearer data on

the use of a particular lexeme from within the noise generated by orthographic and inflectional variation.

Specifically trained models exist for this purpose, including RNNTagger (Schmid 2019), a part-of-speech tagger and lemmatizer, 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.<sup>4</sup> Given the inherent variability in the Middle 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 4, 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: *oïr* 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 5 contains a version of the same stanza tagged and lemmatized using this method. Advantages of this model include the fact that specific details of the word forms, such as verb tense, can be included in the tagging, and the model is more successful at providing consistent lemmata across different forms.<sup>5</sup> This tagging was only achieved after some human feedback, however, since the initial output was less successful, mistakenly claiming, for example, that “voudront” is conditional, and citing “m’orront” as a single word form.<sup>6</sup>

In the case of the first major error, a model trained on manually tagged text would likely never return “voudront” as conditional, unless there were some mistake in the training data. ChatGPT, meanwhile, for which the training data are far more generalized, has hallucinated, in that it has given the

---

**4** 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 lemmatizer, 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.

**5** 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.

**6** We also note the fact that ChatGPT appears not to be able to treat sentence-internal punctuation as tokens. Given that I have not standardized the punctuation in these texts, this is of minimal importance here.

Table 4. Verses 1–4 of *Abesse*, tagged and lemmatized 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        |
| se        | PROper  | soi       |
| traient   | VERcjg  | traire    |
| vers      | PRE     | vers      |
| moy       | PROper  | moi       |
| ,         | PONfbl  | ,         |
| si        | ADVgen  | si        |
| m'        | PROper  | je        |
| orront    | VERcjg  | ouïr      |
| recorder  | VERinf  | recorder  |
| Comment   | ADVsub  | comment   |
| la        | DETdef  | le        |
| mere      | NOMcom  | mere      |
| Dieu      | NOMpro  | Dieu      |
| scet      | VERcjg  | savoir    |
| les       | DETdef  | le        |
| bons      | NOMcom  | bon       |
| conforter | VERinf  | conforter |
| .         | PONfrt  | .         |

Table 5. Verses 1–4 of *Abesse*, tagged and lemmatized 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        |
| se        | PRO:PER       | se        |
| traient   | VER:pres_subj | traire    |
| vers      | PRP           | vers      |
| moy       | PRO:PER       | moi       |
|           | —             |           |
| si        | ADV           | si        |
| m         | PRO:PER       | me        |
| orront    | VER:futur     | ouïr      |
| recorder  | VER:infi      | recorder  |
| Comment   | ADV           | comment   |
| la        | DET:F:S       | le        |
| mere      | ADJ:F:S       | mère      |
| Dieu      | N:M           | dieu      |
| scet      | VER:pres      | savoir    |
| les       | DET:M:PL      | le        |
| bons      | ADJ:M:PL      | bon       |
| conforter | VER:infi      | conforter |
| .         | SENT          | .         |

response that it predicts is most likely to be accepted, without having determined that it is in any way “true.” The second major error, ChatGPT’s inability to recognize “m’orront” as two words, is a simple error of segmentation that suggests an underlying inability to separate words. 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.

Both errors therefore reveal the deficiencies of ChatGPT for the task of lemmatization of Middle French in that, as a natural language model, it has not been trained specifically for the purposes of lemmatization, and so does not necessarily have any underlying conception of the linguistic content, rather the model simply produces an output which it predicts will be accepted as correct (OpenAI 2022b). Given the presence of these errors, as well as the technical difficulties in obtaining reproducible outputs from LLMs at scale, I have therefore ultimately preferred the output of RNNTagger, on the understanding that the tagging may not always be completely accurate.<sup>7</sup> 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).

## 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. Statistical analysis in a vacuum is of little value, however, since noticing that a particular feature is used in a certain way in the *J-corpus* tells us nothing about how common that feature may have been in the work of other authors. To allow for more evidence-based arguments regarding which features of the *J-corpus* may be statistically distinct in Jehan’s work, it is therefore necessary to provide reference corpora against which we may test our hypotheses. The analysis in

---

<sup>7</sup> RNNTagger attempts to provide morphosyntactic tagging for some languages, such as Modern French, but does not do so for Old or Middle 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.

this study makes use of five reference corpora serving a specific comparative purpose, all of which have been lemmatized with RNNTagger to allow word-based and lemma-based analyses.<sup>8</sup>

The *A-corpus* (102,576 words) consists of the surviving poems which most closely resemble those previously attributed to Jehan. Texts have been included if they meet the criteria of being short, narrative medieval French poems written in monorhyme alexandrine quatrains. They must be written in the third person, and with seemingly moralizing, didactic aims. The resulting collection of twenty-six poems therefore offers the closest possible surviving texts to those of the *J-corpus*, allowing us to identify which features may be unique to Jehan, and which may result from the kinds of poems that he chose to write. Sixteen of these texts are saints' lives or miracles associated with a particular saint, four are concerned with a named protagonist or protagonists, and a further six largely treat their characters anonymously.<sup>9</sup> Since the corpus contains all those texts which previous scholars have explicitly rejected as being the work of Jehan, and more besides, the *A-corpus* also enables us to take a more comprehensive view of Jehan's *œuvre*, starting from first principles by allowing the data themselves to reveal whether a given text may merit inclusion.

A second point of comparison, the *F-corpus* (118,761 words), contains texts from an established generic grouping, consisting of 101 *fabliaux* transcribed and encoded as part of a data release by the Laboratoire IHRIM at the ENS de Lyon (Pierreville 2024). The texts of the *F-corpus* are the work of several different authors, and many are anonymous, but the texts all fall within a single genre. By comparison with the *F-corpus*, we are thus able to identify features of the *J-corpus* which may be present only because its texts all fall within the same genre, as opposed to when these features can be said to be more comfortably the result of shared authorship.

---

**8** Short titles for the *J-* and *A-corpora*, as well as the editions consulted when assembling the *A-corpus*, are given in the frontmatter. For the contents of the *F-*, *R-*, and *B-corpora*, which were assembled by others, see the citations given for each. For details of the contents of fr. 24432, forming the *P-corpus*, see Dows-Miller (2025b).

**9** Three of the texts are apparently fragmentary: *Yves* (at end), *Mauvais riche* (at start and end), and *Dieu d'Amours* (at end). I am aware of four other texts which meet these criteria but which remain unedited, namely the alexandrine monorhyme quatrain versions of the lives of sainte Agnes (Carpentras, Bibliothèque Inguimbertaine, MS 106), saint Antoine de Padoue (Paris, Bibliothèque nationale de France, MS français 2198), sainte Catherine d'Alexandrie (Manchester, John Rylands University Library, MS French 6), and saint Eustache (Paris, Bibliothèque nationale de France, MS français 1555).

Our third reference corpus allows us to consider this question from the opposite direction. The *R-corpus* (77,108 words) consists of the complete works of Rutebeuf, as contained within Bastin and 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 *J*-, *A*-, or *F-corpora*, and therefore the *R-corpus* provides evidence of features which may be more strictly explicable by a shared-author hypothesis, as opposed to features which can simply be explained by the fact that the texts of the *J-corpus* share formal and generic similarities.

The *P-corpus* (279,204 words), meanwhile, allows us to take a more codicological perspective, since it consists of the sixty-four texts found within fr. 24432 which are not part of the *J-corpus*, that is, they have not been previously attributed to Jehan.<sup>10</sup> The contents vary in genre, form, style, chronology, and geographic origin, and so unlike the *A-corpus* in particular, the *P-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 alongside most of the *J-corpus*. It is this corpus which therefore allows us to look at fr. 24432 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 fr. 24432, it has sometimes been necessary to subdivide the *P-corpus* to provide the appropriate data for the method of analysis at hand. This sub-corpus, which I term the *P1-corpus* (160,741 words), includes only those fifty-five texts from the *P-corpus* which are consistent in that they are all entirely in verse, and which are not fragmentary. Given that *Guillaume*, *Robert*, and *Vision*, included in the *P-corpus*, also meet the criteria for inclusion in the *A-corpus*, they are shared across both.<sup>11</sup>

Our final 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

**10** The name of the corpus corresponds to the siglum which is often assigned to the whole of fr. 24432 in descriptions of the manuscripts containing *fabliaux* (cf. Noomen and van den Boogaard 1983–1998).

**11** The same is true for several texts in the *F*- and *R-corpora*, but in all cases the copies of these texts in the *P-corpus* have not provided the model for the texts' editions.

of the *Base de français médiéval* (BFM; Guillot-Barbance et al. 2022).<sup>12</sup> 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 shared-author hypothesis for the former. 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 formal, generic, and temporal connections to the *J-corpus*, have been consulted where necessary for a particular direction of analysis.

These five reference corpora and their subsets give us the opportunity to consider the *J-corpus* from several new perspectives. Rather than imposing modern criteria for what counts as suitable evidence for shared authorship, the bespoke methods contained within this chapter will allow us to encounter Jehan's *œuvre* in amongst the enormous diversity of medieval French literature, without seeking to simplify it. In so doing, I present an approach that is widely applicable to other questions of uncertain medieval authorship.

---

**12** RNNTagger was in fact trained on several texts within the *B-corpus* which have been manually lemmatized (Lavrentiev 2023). Rather than use only the manually tagged texts, however, I have chosen to re-tag the whole corpus from plain text, 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.

## Chapter 2

### FORM

**THE OBVIOUS POINT** of departure for our examination of the texts attributed to Jehan de Saint-Quentin is *Chevalier et escuier*, the only poem 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 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. 221–24)

---

**ABSTRACT** In this chapter, we begin our holistic examination of Jehan de Saint-Quentin's *œuvre* by considering the structural and formal characteristics that define the *J-corpus*, taking *Chevalier et escuier*, the only poem containing an act of self-attribution, as our starting point. The analysis of text length reveals that the *J-corpus* is predominantly united by brevity, but also contains a distinct "long-text" sub-group of poems approximately four to five times the median length, to which we can also add *Guillaume*, *Robert*, and *Richard*. To contextualize these findings, I use the *P1-corpus* to assess the *J-corpus* within the specific codicological context of fr. 24432, while the *A-*, *F-*, and *R-corpora* provide broader generic and authorial comparisons, revealing that Jehan's texts display a level of consistency in length that exceeds the work of Rutebeuf, for whom attribution is more secure. Beyond questions of text length, I identify a positive correlation between text length and prologue length, a feature shared with the *A-corpus* but appearing with greater consistency in the work of Jehan.

---

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.<sup>1</sup> 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 several 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*).<sup>2</sup> 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 poems he attributes to Jehan, and which we likewise find in the texts added by subsequent scholarship.

Morawski (1939, 330) observes that these prologues and epilogues in the corpus share 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, particularly Mary, with exhortative structures such as *or prions* and *si prions* especially frequent.<sup>3</sup>

Nevertheless, as Munk Olsen (1963b, 151) notes, there is nothing unusual or surprising about the use of prologues and epilogues in Jehan’s poems,

<sup>1</sup> Previous work on other authors has established that third-person self-reference is not an uncommon phenomenon. See, for example, Marnette (1998).

<sup>2</sup> The length of *Florence* supplied here does not exactly correlate with the length of the text as it appears in fr. 24432, given the absence of folio 223 from the codex, which I assume to be a later loss. Comparison with Grenoble, BM, MS 319, the other manuscript in which *Florence* appears confirms that the lacuna consists of seventy-two lines of verse, corresponding to a single folio containing four thirty-six-line columns, with each alexandrine line split into two hexasyllable hemistichs.

<sup>3</sup> For more on this feature, see the discussion of terms of address in Chapter 4.

since such sections are commonly found in other texts with the same verse 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.

## 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

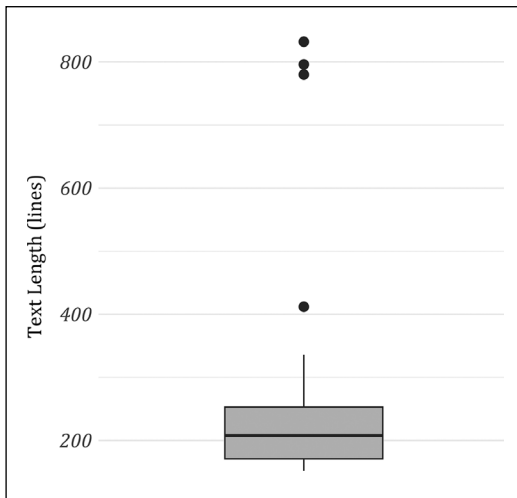


Figure 2. Box-and-whisker plot of the lengths of the *J-corpus*.

the box-and-whisker plot in Figure 2, we find that the poems have a median length of 208 lines, and eighteen of the twenty-four texts (75%) come in at under three hundred lines in length. The corpus is therefore mostly united in its brevity, but the plot also reveals the significant gap in length between three outlying, longer texts (*Florence*, *Anelés*, and *Buef*), and the next longest text (*Chanoines*), which is itself still an outlier despite being nearly half the length of the longest at only 412 lines.<sup>4</sup>

These three longer texts all appear relatively close together in fr. 24432, 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.<sup>5</sup>

<sup>4</sup> Our definition of an outlier here, as is common practice, is a value which is less than 1.5 times the interquartile range below the first quartile (25% of the data), or greater than 1.5 times the interquartile range above the third quartile (75% of the data).

<sup>5</sup> *Florence* appears on fols. 215rb–226va, *Anelés* on fols. 231ra–241rb, and *Buef* on fols. 247ra–257va.

This statistical evidence, combined with the codicological evidence, suggests that we may see the *J-corpus* as dividing into two groups, those around the two hundred-line mark which are written with brevity and concision in mind, perhaps for brief recitation, and those around eight hundred lines in length, which provide space for a more expansive narrative. Certainly, this alone does not suggest that the longer poems were not composed by Jehan,

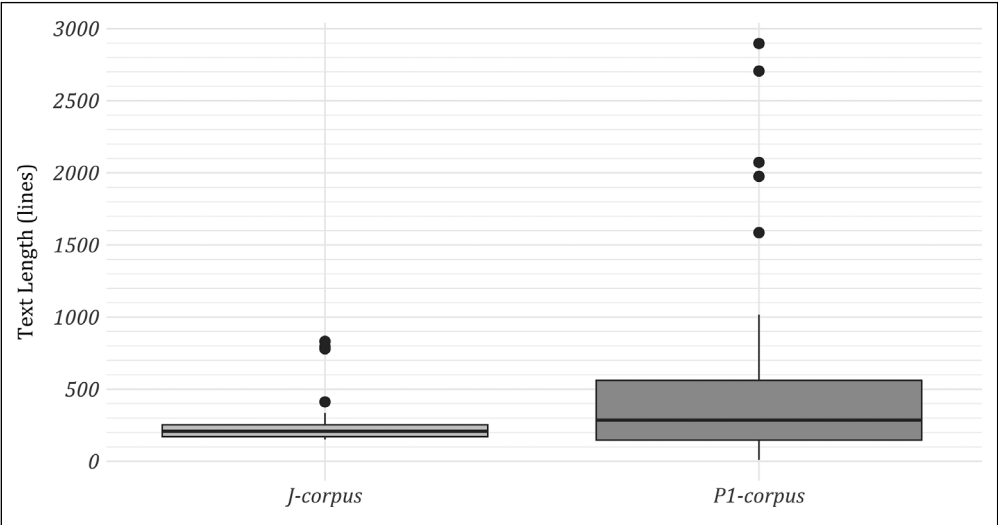


Figure 3. Box-and-whisker plot of the lengths of the texts of the *J-corpus*, compared to those of the *P1-corpus*.

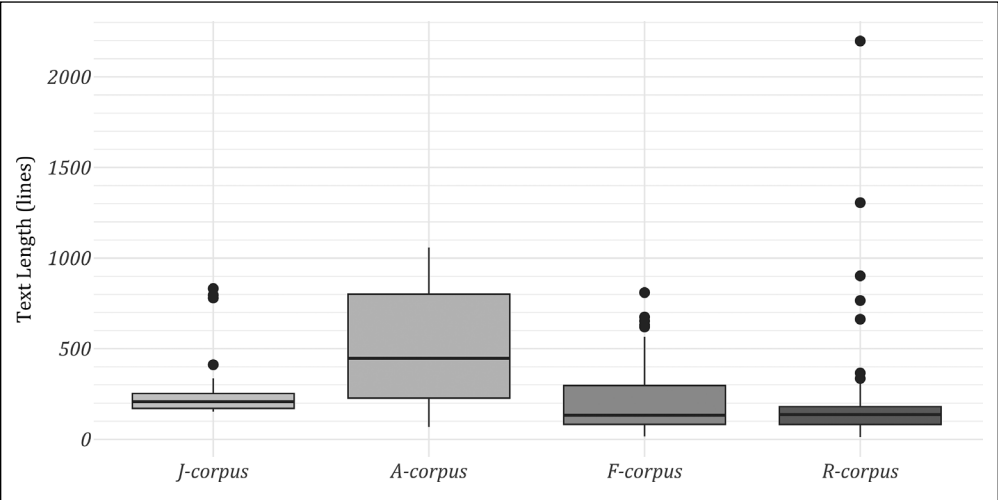


Figure 4. Box-and-whisker plot of the lengths of the texts of the *J-corpus*, compared to those of the *A*-, *F*-, and *R-corpora*.

but it does validate Moraswki's decision to view *Florence* and *Anelés* as a pair distinct from the rest of the corpus, to which we may also add *Buef*.

Fruitful examination of text length comes also from comparison of the *J-corpus* to those non-fragmentary poems found elsewhere in fr. 24432, contained in the *P1-corpus*, which can help us identify how usual the length of Jehan's texts is for the whole codex. As presented in Figure 3, 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.<sup>6</sup> Further, the longest texts of the *P1-corpus*, of which the most extreme example is the *Dit du tournoiment d'Antecrist et d'euls de son barnage* at a length of 2,897 lines,<sup>7</sup> reveal that *Florence*, *Anelés*, and *Buef*, while long by comparison to the other texts of the *J-corpus*, are not especially long within the codex as a whole.

However, while the *P1-corpus* is useful for contextualizing the *J-corpus* in the manuscript in which most of its texts appear, it does not allow like-for-like comparison. To do this, we must consider the texts of our more cohesive reference corpora, namely the *A-corpus* of narrative poems written in alexandrine monorhyme quatrains, the *F-corpus* of *fabliaux*, and the *R-corpus* of Rutebeuf's works.

Figure 4 displays box-and-whisker plots for the *J-corpus* alongside these other corpora, which reinforce our suggestion that Jehan's *œuvre* may in fact be notably limited in the variation of text length. Firstly, while the median text lengths for the corpora are roughly comparable between the *J*-, *F*-, and *R-corpora* (*J*: 208; *F*: 176.5; *R*: 137), the *J-corpus* has the narrowest range between the first and third quartiles (*J*: 82; *F*: 244.75; *R*: 98), suggesting a level of consistency which exceeds even that of the *R-corpus* for which the attribution of authorship is more assured.

The *A-corpus*, by contrast with these other corpora, has both a much higher median text length than the *J-corpus* (448) and a much wider inter-quartile range (573), when we exclude the three fragmentary texts. Notably, of the twenty-three non-fragmentary texts of our *A-corpus*, only seven are within the range of the *J-corpus* and would not be classed as outliers within it, namely: *Roy amie* (200), *Quatre sereurs* (204), *Leu* (252), *Sebastien* (256), *Basile* (352), *Jehan Paulus* (370), and *Confesseur* (374). Variation in text

<sup>6</sup> Note that the discussion here is of lines of verse for a given text, not orthographic lines of text as they appear in the manuscript. Alexandrine verse in fr. 24432 is usually broken into two hemistichs by the scribes, whereas hexasyllable, octosyllable, and decasyllable are usually presented on a single line.

<sup>7</sup> For an edition, see Tarbé (1851).

length up to about one thousand lines therefore appears particularly common for narrative poems in alexandrine monorhyme quatrains, making the appearance of a narrower, more coherent group within the *J-corpus* even more significant.

Taking a closer look at the texts of the *A-corpus* which I argue to be the work of Jehan, namely *Robert*, *Guillaume*, *Richard*, and *Jehan Paulus*, we find that *Robert*, at some 1,016 lines, becomes the longest of Jehan's poems, followed closely by *Guillaume* (948 lines) and *Richard* (818 lines). All three texts do therefore fit reasonably comfortably alongside *Florence* (832), *Anelés* (780), and *Buef* (796) within the longer grouping of texts from the *J-corpus*, especially given that *Robert* immediately precedes *Florence* in fr. 24432. *Jehan Paulus*, meanwhile, at 370 lines in length, sits just above the third quartile of the *J-corpus*, and would not be considered an outlier in the data.

Based on these data, then, we are already able to identify two core sub-groups within Jehan's works, one containing poems of around 200 lines in length, and another, which we may term the "long-text" sub-group, containing poems roughly four to five times longer. Based on the outliers that appear in the *R-corpus* of Rutebeuf's poems, of which the most extreme is *La Vie sainte Elysabel*, at 2,198 lines, we note that this division does not suggest that one group ought to be excluded from Jehan's *œuvre*. An author is not a one-trick pony, and may be capable of modifying the form of his poems where he deems it necessary. 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 in the *F-corpus*.

## Prologues and Epilogues

Taking a similar approach to the prologues and epilogues of the *J-corpus*, we observe that *Chevalier et escuier* is unique within the corpus in that its epilogue is longer than its prologue, a feature not seen in the other texts.<sup>8</sup> 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 epilogue reaches this length, rather than suggesting that the

---

<sup>8</sup> *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).

whole process of attribution is mis-aligned. The intensity of the epilogue's devotional content offers a suggestion for why Jehan extended this section, choosing to name himself to highlight his own piety, in a way that he did not do elsewhere.

Examining the non-fragmentary texts of the *A-corpus*, we find a greater degree of variation, with fifteen texts displaying a longer prologue than epilogue, as in most of the *J-corpus*,<sup>9</sup> three where the opposite is true (*Leu*, *Sebastien*, *Quatre sereurs*), and four which either do not have a prologue (*Servais*, *Vieillards*, *Roy amie*) or an epilogue (*Thomas*). It is therefore clear that this feature, again, increases our sense of the cohesion of the *J-corpus*, and we can build on this by taking a more statistical approach the lengths of the prologues and epilogues.

Prologues are usually short in the *J-corpus*, with the two longest prologues (eighty-four and forty lines) appearing in two of the longest texts, *Anelés* and *Buef* (780 and 796 lines), both of which use the prologues to offer a summary of the plot to come. Despite some notable anomalies, such as the prologue of thirty-two lines found in *Beguine*, a text of only 204 lines in length, 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. The texts from within the *A-corpus* which we are including Jehan's *œuvre* also fit this observation, with the three longer texts *Robert*, *Guillaume*, and *Richard* displaying prologue lengths of nineteen, thirty-six, and sixteen respectively, all within the range of the "long-text" grouping within the *J-corpus*, while *Jehan Paulus* has an eight-line prologue and thus sits comfortably alongside the remainder of the *J-corpus*.

The positive correlation between text length and prologue length is not a feature unique to the *J-corpus*, however. Across all the texts of the *A-corpus*, we observe a correlation coefficient between text length and prologue length of 0.48537 ( $p = 0.01889$ ), suggesting that there is an increased desire across most longer poems in written alexandrine quatrains to engross an audience in the story before beginning the narrative, in a way that is not necessary in shorter poems. This hypothesis would also go some way to explain why

<sup>9</sup> Barbe, Basile, Cristofle, Dieudonnée, Grégoire, Jehan Ewangeliste, Jehan Paulus, Confesseur, Vision, Thibaut, Guillaume, Robert, Richard, Troyes, and Dés.

the longer prologues in the *J*- and *A*-corpora often have a greater focus on narrative content than the shorter ones, rather than dwelling on any didactic message, since this enables the author to better prepare an audience for what is to come.

However, beyond the observation that a longer text is generally accompanied by a longer prologue, we should note a point of difference between the two corpora, namely that the *A*-corpus as a whole displays a greater standard deviation ( $\sigma$ ) in the lengths of the prologues, both in real terms and as a percentage of the mean (*J*:  $\sigma = 16.2$ , 93.3% of the mean; *A*:  $\sigma = 35$ , 160.2%). 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, but Jehan appears to do this in a more consistent way than the various authors of the *A*-corpus. Again, this reveals greater cohesion between the texts of the *J*-corpus than between those of the *A*-corpus, broadly reinforcing our shared-author hypothesis.

While there are indicators of distinctiveness in the use of prologues between the *J*- and *A*-corpora, the trends followed by the epilogues in the *J*-corpus appear as much determined by formal, generic considerations as they are by authorial ones. Evidence for this comes from the fact that for both corpora there is no statistically significant link between total text length and epilogue length (*J*:  $p = 0.2608$ ; *A*:  $p = 0.0863$ ), or prologue length and epilogue length (*J*:  $p = 0.568$ ; *A*:  $p = 0.8341$ ). 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 or form, 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 a combination of authorial preference and 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 (*J*:  $\sigma = 3.50$ , 54.2% of the mean; *A*:  $\sigma = 7.67$ , 102.6%) is also a possible indication that authors of both corpora 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*.<sup>10</sup>

Our examination of the formal features of prologues and epilogues has therefore revealed an apparently greater authorial influence in prologue length than epilogue length, which intersects with broader trends that apply across both the *J-* and *A-corpora*. We observe that, with limited, explicable exceptions, there is a coherence to the texts of the *J-corpus* that is not found to the same extent in the *A-corpus*. We may therefore tentatively take these data as initial evidence of the shared-author hypothesis for the *J-corpus*, while gaining a clearer understanding of the generic framework in which Jehan operated, a question which we will return to in the next section.

As we have already observed, as we move through this study, there will be no single smoking gun that definitively proves or disproves the attribution of any text to Jehan. Rather, it is the holistic combination of various such unifying features that we have now begun to build, which 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.

---

**10** The epilogues of *Robert*, *Guillaume*, *Richard* and *Jehan Paulus* have epilogues of one, one, three, and four lines respectively, aligning with our observation that epilogue length in both the *J-* and *A-corpora* shares little to no statistical relation to the length of the rest of the text, but that epilogues in Jehan's works are almost always shorter than prologues. In the case of *Jehan Paulus*, the text's use of an eight-line prologue and four-line epilogue lines up precisely with those of *Pommes* and *Povre Chevalier* within the *J-corpus* (although the latter may well once have had a twelve-line prologue, given that lines 1–2 do not rhyme with lines 3–4).

## Building the Case for the Shared-Author Hypothesis

The formal and structural analysis carried out in this chapter provides the first tranche of evidence to support the claim that Jehan's previously attributed *œuvre* can be reinforced and expanded. It reveals that the *J-corpus* is split into two broadly coherent sub-groups by length: a collection of shorter texts running to around two hundred lines, and a "long-text" sub-group containing *Florence*, *Anelés*, and *Buef. Guillaume*, *Robert*, and *Richard* align with the latter group, while *Jehan Paulus* is associated with the former, reinforcing the association of all four texts with Jehan. Further, all our additional texts align with the observation that Jehan's longer poems generally have longer prologues, while the same is not true for epilogues. This formal consistency, alongside the codicological evidence provided by *Robert's* proximity to *Florence* in fr. 24432, aligns these four texts with the *J-corpus*, and builds the case for the shared-author hypothesis.

## Methodological Contribution

In this chapter, I have demonstrated the utility of macro-level formal and structural analysis when applied to a corpus with significant internal variation. By employing data-driven visualizations, we have moved beyond a reliance on simple averages to identify how outliers can reveal distinct textual groupings within the work of an individual author. Further, the application of statistical correlation tests provides a reproducible, falsifiable framework for examining the internal formal and structural consistency of a poet's output. Our approach has revealed that structural features like prologue length, while often comparable to works of a similar form by other authors, cannot solely be explained by generic requirements, and are instead often the result of authorial choices that can be quantified and compared.

## Chapter 3

### 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. As discussed in the previous chapter, this is particularly noticeable in the prologue-epilogue pairs, but it is found throughout the narrative as well, and 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 hérissier les cheveux sur la tête" (1939, 345–46),<sup>1</sup> which again, as Munk Olsen (1963b, 152) notes, is in no way astonishing for the didactic material of the period, and many similar tales appear in the *A-corpus*, yet it does add a further sense of thematic unity to the *J-corpus*. In the case of *Chevalier et escuier*, while its contents may not be as shocking as the infanticide of *Romme*, *Pecherresse*, or *Enfant rosti*, the two companions' initial willingness to renounce God and the saints, as well as the squire's gruesome ending at the bottom of a hangman's rope, is certainly intended to shock, while the knight's redemption provides a contrasting example of good moral behaviour.

---

<sup>1</sup> "strange or horrible subjects, apt to make one's hair stand on end."

---

**ABSTRACT** Moving beyond questions of structure, this chapter examines the use of generic descriptors in the *J-corpus* and our reference corpora, considering both internal self-descriptors and external paratextual labels. Comparing the *J-corpus* to the *A*-, *P*-, and *B*-corpora, I examine the distribution and collocations of three primary descriptors (*miracle*, *exemple*, and *dit*), and demonstrate that Jehan's use of these terms stands out amongst our corpora. Specifically, I identify an apparent preference for the phrase (*un moult tres*) *bel exemple*, and the frequent association of generic terms with verbs of orality, which do not appear with such frequency elsewhere. Finally, I identify that the ubiquitous use of the term *dit* by modern scholarship to describe Jehan's poems stems largely from the choices made by the rubricator(s) of fr. 24432, rather than from any authorial preference, but that the consistent use of these markers acts as further evidence in support of the shared-author hypothesis.

---

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. Such *post rem* generic classification has long been acknowledged to be lacking,<sup>2</sup> given that modern generic descriptors so often fail to encapsulate medieval conceptions of genre. More recently, a more holistic and historically informed understanding of genre has been offered through the "poétique des genres"<sup>3</sup> enacted by Gingras (2016, 8), encompassing "aussi bien la forme revendiquée par l'auteur que le genre attribué par l'histoire de sa réception, depuis les rubriques des copistes et les associations suggérées par la mise en recueil, jusqu'aux classements canoniques établis par la critique depuis la fin du XIX<sup>e</sup> siècle."<sup>4</sup>

However, it is of minimal significance to this study exactly which generic descriptor modern scholarship should choose to use for Jehan's poetry. The purpose of our analysis here is neither to prescribe nor to proscribe any specific generic term as a suitable descriptor for his texts. Nevertheless, questions of genre do still have much to offer the study of Jehan's *œuvre* from the perspective of the attribution of authorship. This is especially true when we align our approach with Gingras's genre poetics, considering the self-descriptors contained within and around Jehan's poems in their manuscript contexts, as well as asking why modern scholarship has chosen to describe the texts in the way that it has.

Taking this approach can help us in our search for unity within the *J-corpus*, and help us better to understand how the texts were conceptualized by their author, as well as their earliest audiences. In particular, what is most meaningful for our purposes here 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 poems, then this can contribute to the corpus' sense of coherence, while suggesting that the scribes who copied them may also have been aware of this unity. Neither of these features is enough to guarantee shared authorship on its own, yet they contribute to our growing, holistic picture of which elements of the *J-corpus* form a unified whole, assisting in the process of attribution.

---

**2** For discussion of such challenges see, for example, Jauss (1970), Zumthor (1972), and Lacy (1987).

**3** "genre poetics."

**4** "both the form claimed by the author and the genre attributed by the history of its reception, from the rubrics used by copyists and the associations suggested by the compilation, to the canonical classifications set out by critics since the end of the 19<sup>th</sup> century."

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 in turn, comparing their usage with what we might expect based on an analysis of our broader reference corpora. Following this, we look beyond the poems themselves, turning to an examination of the use of generic descriptors in fr. 24432's paratexts.

### **miracle**

The self-descriptor *miracle* appears on several occasions in the *J-corpus*. As a generic 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* 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* is found in fr. 24432 on fols. 315va–316vb, immediately following Jehan's *Beguine*, perhaps suggesting an acknowledgement of similarity between these texts on the part of the manuscript's compiler.<sup>5</sup>

The similarities between de Coincy's *miracle* and *Respon*, another text of the *J-corpus*, are stark. In Jehan's 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. 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, convincing him to make good on his promise and embrace a life of chastity. Despite differing metres, the epilogues of both texts are notable in their similarity.<sup>6</sup>

<sup>5</sup> The text in question is the catchily titled *Le dit de l'enfant qui mist l'anel que s'amie li avoit donné ou doit a l'ymage de Nostre Dame*. For an edition, see Koenig (1955–1970, ii: 197–204).

<sup>6</sup> The surviving copies of Gautier de Coincy's text vary in their formulation of the epilogue, and the version found in fr. 24432 is not noted by Koenig (1955–1970) in

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. 189–92)

Or prion Dieu le fil Marie,  
 Qu'il nous doint a touz bone vie  
 Et apres vie pardurable,  
 Car celui est bon heritage.

(fr. 24432, fol. 316vb, lines 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 this leads Léonard (1996, 327) to argue that many of Jehan's texts are "incontestablement des *miracles*."<sup>7</sup> And yet, this is not straightforwardly true of all of the texts in the *J-corpus*, including *Beguine*, the text that immediately precedes de Coincy's *miracle* in fr. 24432. This text falls into Kleist's (1973, 66–67) category of "negativen Mirakeln," that is, "negative miracles" or stories containing miraculous content but in which a sinner is damned rather than saved. Within *Beguine*, the eponymous *religieuse* falls into league with the Devil and places the Host into a box with a toad, prompting a Eucharistic miracle in which the Host cries out like a newborn 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 positive ending for the sinner, but it is nevertheless miraculous in its content.<sup>8</sup>

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. As we have noted, *Chevalier et esquier*, the text containing the instance of Jehan's self-attribution 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 we should refer to the texts of the *J-corpus* as *miracles*. Medieval genre is of course

---

any of the ten manuscripts consulted for his edition. I have not found a comparable formulation elsewhere.

7 "undeniably miracles."

8 Tales resembling this one, we may note, are common also in *exempla*, as discussed on pp. 46–53. See, for example, the Eucharistic tales included in the *Index exemplorum* (Tubach 1969).

a thorny question, 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 categorized many of his texts as *miracles*, had he chosen to do so.

From the perspective of authorial attribution, a digital textual analysis of the *J-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 *J-corpus*, we find eighteen uses of *miracle*, all of which refer to a greater or lesser extent to an individual miracle within the narrative, but three of which can also be said to function as generic descriptors which refer to the whole text. These are found in *Narbonne* and *Pecherresse* (twice), and are displayed in Table 6.

Table 6. Uses of *miracle* as a generic descriptor within the *J-corpus*.

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

Given that, as we have seen, the poems of the *J-corpus* share numerous 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 that had any generic implications. This observation also appears to be true of the *A-corpus*, in which there are some thirty-eight uses of the term *miracle*—an unsurprisingly higher rate of use given more of the texts are explicitly hagiographic—of which only one could also be read as a generic descriptor, as in Table 7.

Table 7. The single use of *miracle* as a generic descriptor within the *A-corpus*.

| Text           | Line | LH                           | KW      | RH      |
|----------------|------|------------------------------|---------|---------|
| <i>Thibaut</i> | 1056 | En cestui present livre cest | miracle | escrist |

The sense that *miracle* in both the *J-* and *A-corpora* appears as a narrative descriptor, rather than a generic one, is reinforced by the fact that *miracle* in either use is often accompanied by *bel*, with this term characterizing the

moral, rather than aesthetic, beauty of the narrative itself. For the *J-corpus*, in addition to the two generic examples seen above, there are a further four narrative cases where a form of *bel* appears within two tokens either side of *miracle*, shown in Table 8, such that a third of the uses of *miracle* found anywhere in the *J-corpus* are accompanied by this word. These examples once again give us cause to consider *Florence* and *Buef* as a pair, which may be notably distinct from other texts of the *J-corpus*.

Table 8. Other uses of *miracle* alongside *bel* within the *J-corpus*.

| Text            | Line | LH                             | KW      | RH                |
|-----------------|------|--------------------------------|---------|-------------------|
| <i>Romme</i>    | 185  | La fist moult <b>biau</b>      | miracle | Nostre Dame Marie |
| <i>Narbonne</i> | 80   | Or oez <b>biau</b>             | miracle | qui oïr leouldra  |
| <i>Florence</i> | 803  | Sailli hors sanz main metre le | miracle | fu <b>bel</b>     |
| <i>Buef</i>     | 644  | Vous porrez ja oïr un          | miracle | moult <b>bel</b>  |

Similarly in the *A-corpus*, seven of the thirty-eight uses of *miracle* are accompanied by *bel*, as in Table 9, where we note the appearance of such a structure in *Richard*, albeit not as a generic descriptor.<sup>9</sup> The difference in rates between the *J-* and *A-corpora* is not statistically significant, suggesting that the use of *bel* alongside *miracle* in the *J-corpus* is hardly unique, but there are notable differences in the ways that *bel* is used in both. In the *J-corpus*, *bel* appears to be preferred at the rhyme, as in *Florence* and *Buef*, while its variant *biau* is consistently used line-internally and appears like an epithet, perhaps included for reasons of metre. In the *A-corpus*, meanwhile, *bel* and *biau* only appear immediately before *miracle*.

Table 9. Uses of *miracle* alongside *bel* within the *A-corpus*.

| Text                     | Line | LH                      | KW       | RH                          |
|--------------------------|------|-------------------------|----------|-----------------------------|
| <i>Alexis</i>            | 300  | Ja orez <b>biau</b>     | miracle  | et plain de grant pitiés    |
| <i>Basile</i>            | 284  | Et des <b>bielles</b>   | miracles | ne me repuis plus taire     |
| <i>Cristofle</i>         | 708  | Onques plus <b>beau</b> | miracle  | n'oïstes maiz compter       |
| <i>Dieudonnée</i>        | 253  | La moustra <b>biaus</b> | miracles | Dieus, qui set tout et voit |
| <i>Jehan Ewangeliste</i> | 537  | Certes c'est <b>bel</b> | miracle  | cil doivent savoir tuit     |
| <i>Leu</i>               | 100  | Si ourrés <b>beau</b>   | miracle  | qui est moult avenant       |
| <i>Richard</i>           | 315  | Ce <b>beau</b>          | miracle  | cy si advent a Rouen        |

<sup>9</sup> In the title of its incunable edition, the text is described as a “romant.” A narrative use of *miracle* is also observed within *Jehan Paulus* (v. 209).

Further, when comparing the use of *bel* to that found our more general reference corpora, namely the *P-corpus* of the texts in fr. 24432 not attributed to Jehan, and the large *B-corpus* which contains texts from a wide variety of generic, diachronic, and diatopic contexts, we find that the use of the term alongside *miracle* is unusually high in both the *J-* and *A-corpora*, but especially the *J-corpus*, when taken in this broader context. Both display statistically significant differences with the *P-corpus*, in which *bel* does not appear anywhere alongside the twenty-one instances of *miracle* (*J*:  $p = 0.00593$ ; *A*:  $p = 0.04327$ ) but statistical significance only applies for the *J-corpus* relative to the *B-corpus*, for which the figure is 10.24% (eighty-one of 791 uses of *miracle*; *J*:  $p = 0.00434$ ; *A*:  $p = 0.10713$ ). This means that, while *bel* appears to be a term used alongside *miracle* with high frequency in narrative poems written in alexandrine monorhyme quatrains, it is especially common in the texts of the *J-corpus*.

A further point of contrast between the *J-* and *A-corpora* comes in the fact that, in the *A-corpus*, we find occasional use of other descriptors alongside *miracle* which do not appear in the *J-corpus*, particularly *haut* and *apert*, as in Table 10. Both terms are focused on the divine significance of the miraculous events described, and it is therefore unsurprising that all these examples appear in texts which are explicitly saints’ lives. This, we may remind ourselves, is only true for the *J-corpus* in the case of *Sauveur*, and these examples thus reinforce our sense of Jehan’s apparent preferred subject matter: miraculous, holy events, which are mobilized in support of a didactic goal, rather than the veneration of an individual saint.

Table 10. Uses of additional descriptors alongside *miracle* within the *A-corpus*.

| Text                     | Line | LH                         | KW       | RH                       |
|--------------------------|------|----------------------------|----------|--------------------------|
| <i>Basile</i>            | 144  | Mais les <b>hautes</b>     | miracles | raconter vous volons     |
| <i>Jehan Ewangeliste</i> | 370  | Par tens te monsteroie une | miracle  | <b>aperte</b>            |
| <i>Jehan Ewangeliste</i> | 845  | Oiés tres <b>haut</b>      | miracle  | et bien divers afaire    |
| <i>Jehan Ewangeliste</i> | 887  | Cascun jor i avint une     | miracle  | <b>aperte</b>            |
| <i>Thibaut</i>           | 978  | Que Dés <b>apers</b>       | miracles | par lo cors seint faroit |

Of final note is the fact that *miracle* in the *J-* and *A-corpora* is often associated with verbs of speech and hearing, with seven instances (38.89%) in the *J-corpus* and sixteen (42.11%) in the *A-corpus* in which *miracle* is accompanied by one or more instance of *oïr*, *dire*, *escouter*, *entendre*, (*ra*)*conter*, or *taire*, with *oïr* the dominant form in both corpora. The examples from the *J-corpus* are given in Table 11. This once again contrasts in a statistically sig-

nificant way with the *P-corpus*, which contains only one such example (within Gautier de Coincy’s aforementioned *miracle*), and thus has a rate of such use at 4.76% (one of twenty-one instances,  $p = 0.00953$ ), whereas that of the *B-corpus* is only slightly higher at 8.34% (sixty-six of 791,  $p = 0.00022$ ). The high use of such verbs alongside *miracle* therefore appears to be a particular feature of the kinds of narrative alexandrine poems contained within the *J-* and *A-corpora*, rather than resulting from an individual authorial preference.

Table 11. Uses of *miracle* alongside verbs of speaking and hearing within the *J-corpus*.

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

Through comparison with our reference corpora, the examination of *miracle* therefore reveals that, in a broad sense, Jehan’s use of the term is comparable to that found in other poems in alexandrine monorhyme quatrains. However, there are also cases where the poet associates *miracle* with certain words and actions more than other authors, which gives the texts of the *J-corpus* a sense of unity, and adds to our growing sense of Jehan’s personal style. Ultimately, while he may use the term to talk about miraculous narrative events, Jehan does not appear particularly interested in calling his texts *miracles*.

exemple

Turning to the use of the terms *exemple* and *exemplaire*, we find that, despite several appearances of these descriptors within the texts, there is limited evidence that would qualify the texts of the *J-corpus* for description as *exempla*. Such texts, generally defined as “un récit bref donné comme véridique et destiné à être inséré dans un discours, en général un sermon, pour convaincre un auditoire par une leçon salutaire” (Polo de Beaulieu 2019, 39),<sup>10</sup> saw a

**10** “a brief story presented as true and intended to be inserted into a speech, usually a sermon, to convince an audience with a salutary lesson.”

huge rise in popularity following the Fourth Lateran Council in 1215 (Welter 1927, 1). Their particular success between the middle of the thirteenth and fourteenth centuries, as Polo de Beaulieu et al. (2026) have argued, reflects a “confiance aveugle dans la puissance de la narration,”<sup>11</sup> with preachers using narrative tales within sermons to respond to the challenges of an increasingly urban society, in which “[i]l fallait convaincre un public urbain, qui en majorité ne savait pas lire, et qui était attiré par des croyances jugées hérétiques par l’Église.”<sup>12</sup> To achieve this end, these “historiettes,” these little stories and anecdotes (Berlioz 1994, 235), served a double purpose, both as tales to teach and tales to entertain, with their more colourful elements engaging an audience so as to ensure that they would also pay attention to the moral message at its core. Ultimately, as Taylor notes (2000, 741), it was this focus on entertainment which led to increasing calls “for their elimination or radical reduction” from the fourteenth century onwards.<sup>13</sup>

On the face of it, Jehan’s texts overlap significantly with the canonical understanding of the *exemplum*. The poems of the *J-corpus* all seek both to entertain an audience, while also imparting a moral message. They are also all relatively brief, and while *exempla* were usually not performed in verse, this is not unheard of (Kienzle, 148). And yet, Munk Olsen (1963a, 10) does appear somewhat justified when he argues that Jehan’s texts cannot be described as *exempla*. He notes 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 *Dits* in our collection are not without resemblance to *exempla* [...] The most important difference is that the narrative element, greatly reduced in sermons so as not to exceed the didactic framework, has been expanded and dominates in the *Dits*; but both serve to demonstrate a moral or religious thesis which is presented in their introduction.

11 “blind confidence in the power of narration.”

12 “[i]t was necessary to convince an urban audience, who for the most part could not read, and who were attracted by beliefs deemed heretical by the Church.”

13 On the *exemplum*, see also Brémond, Le Goff, and Schmitt (1982) and Smirnova (2022). For a statistical textual analysis of a corpus of *exempla*, not dissimilar to what we are conducting here, see Berlioz and Guerreau (2019).

It is possible to argue against Munk Olsen's suggestion that *exempla* were inevitably reduced in their narrative focus compared to Jehan's poems. Many *exempla* found in the extant collections are notably abbreviated, as for example in the short vernacular tales of the *Ci nous dit* which contain very little extraneous narrative content,<sup>14</sup> but this does not mean that the same was true for the spoken sermons in which such *exempla* may have been used. The highly condensed written *exempla* which survive, both in Latin and in the vernacular, were seemingly little more than *aide-mémoires* of a broader tale which could then be written up into full stories within a sermon. Taylor (2000, 741) notes that, in France in the later Middle Ages, "[a]lthough prescriptions suggested the maximum length of an hour, Lenten sermons could often extend to two or three hours," and thus *exempla* as they appeared in sermons could certainly have been of a comparable length and narrative complexity to the shorter of Jehan's poems. If, as Munk Olsen himself suggests,<sup>15</sup> Jehan also in some cases developed his tales from sources such as the *Ci nous dit*, then his work could well be much closer to that of sermon writers than Munk Olsen acknowledges.

However, Munk Olsen's point does stand when we consider the didactic framework surrounding the tales. The primary difference between Jehan's poems and *exempla* is primarily in the nature and scale of the didactic context. Kleist (1973, 67) supports this suggestion, arguing that *exempla* "dient als Illustration einer These, deren Wahrheit in der Predigt schon durch Gründe und Argumente bewiesen worden ist."<sup>16</sup> Jehan's texts, meanwhile, differ from *exempla* in that the tales themselves function as the proof of the moral message, rather than an illustration of a message that has already been proven.

Kienzle (2000, 145) offers a theoretical perspective which sheds further light on the matter. Applying Jauss's theory of genres (1970), she distinguishes between "genres with an independent function and those with a dependent function." The *exemplum* is an example of a genre with a dependent function, since it is "encompassed by" a genre with an independent function, namely the sermon. The *exemplum* appears in a sermon only after a lengthy exposition of a particular point of morality, liturgy, or doctrine,

**14** For an edition of these texts, see Blangez (1979–1986), and for recent discussion, see Gingras (2022).

**15** See, for example, Munk Olsen (1978, cxviii), and the discussion of sources in Chapter 8.

**16** "serve as the illustration of a thesis, the truth of which has already been proven in the sermon through reasoning and argumentation."

and serves to reinforce that point, rather than serving independent aims.<sup>17</sup> Jehan's texts, by contrast, carry an independent function, since there is no such accompanying co-text, and his moral messages appear primarily in the prologue and epilogue to a text, where they 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. 161–64)

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 and independently developing his moral argument, rather than explicitly expounding an external argument in a way that may be done with an *exemplum*.

Considering this evidence, the overwhelming conclusion we must draw is that, while Jehan's poems certainly exist in close generic proximity to *exempla*, there is good reason not to describe them as such. As a result, 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 thirty-seven appearances in various spellings of either *exemple* or *exemplaire*, of which nineteen are generic descriptors, as in Table 12, with the others either referring to moral lessons given to characters within the narrative, or referring to a specific element of that narrative.

The use of *exemple* as a self-descriptor therefore appears more consistent across the *J-corpus* than for *miracle*, with appearances in thirteen of the corpus' twenty-four texts. On a surface level, then, the significant frequency with which the term is used in the *J-corpus* would contradict our arguments above in which we dismissed the possibility that Jehan's texts could ever

**17** On the structure of the medieval sermon, see Wenzel (2016).

Table 12. Uses of *exemple* and *exemplaire* as generic descriptors within the *J-corpus*.

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

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, the Latin term does not appear in the *A-corpus* or in any of the texts of the much larger *B-corpus*, and this lack of alternation between the Latin and French forms therefore reveals a general truth that Jehan’s conception of *exemple* and *exemplaire*, and that of other vernacular writers, was 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 moralizing and didactic focus.<sup>18</sup> It is this that can explain

<sup>18</sup> Léonard (1996, 297) observes a similar trend within her larger corpus of *dits*, arguing that *exemple* and *exemplaire* are used in the vernacular to refer to “un récit de forme littéraire, souvent versifiée mais pas toujours” (“a tale in literary form,

the prevalence of *exemple* and *exemplaire* in the *J-corpus*, despite the generic distance 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 distinguish the *J-corpus* from others, we find that of the twelve appearances of the term in the *A-corpus*, the term is used as a generic descriptor on only three occasions, across only two of the corpus' twenty-six texts, as given in Table 13.

Table 13. Uses of *exemple* as a generic descriptor within the *A-corpus*.

| Text                  | Line | LH       | KW      | RH                                        |
|-----------------------|------|----------|---------|-------------------------------------------|
| <i>Leu</i>            | 233  | Par cest | exemple | si creez tres bonne gent                  |
| <i>Quatre sereurs</i> | 1    | Par      | exemple | vorai parler de l'Escripture              |
| <i>Quatre sereurs</i> | 4    | Qu'      | exemple | doit on panre quant la lettre est obscure |

Similarly, from the perspective of the cohesion of the *J-corpus*, and thus the validity of the shared-author hypothesis, we find that, looking across all thirty-seven appearances of the word, there are several 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 *A*-, *P*-, or *B-corpora*. Meanwhile, the shorter construction *moult bel* appears a total of seven times immediately before *exemple* in the *J-corpus* with a rate of 18.92%, which again does not appear in the *A-corpus*, nor in any of the seventy-eight uses of *exemple* in the *P-corpus*.<sup>19</sup> In the *B-corpus*, meanwhile, there is only a single instance of *moult bel* before *exemple* out of the 789 uses of the word (0.13%). Looking at the use of *bel* within any structure containing *exemple*, we observe a rate of 27.03% for the *J-corpus* (ten of thirty-seven instances), 0% for the *A-corpus* (zero of twelve), 7.69% for the *P-corpus* (six of seventy-eight), and 1.01% for the *B-corpus* (eight of 789). There is thus clear statistical significance for the difference in the use of *bel* alongside *exemple* in the *J-corpus* compared to the *A*- ( $p = 0.04456$ ), *P*- ( $p = 0.00572$ ), and *B-corpora* ( $p = 0$ ), such that it appears reasonable to conclude 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.<sup>20</sup>

---

often but not always in verse”), which nevertheless has the same didactic intentions as an *exemplum*.

**19** There are, however, two instances of *tres bel* appearing before *exemplaire*.

**20** Namely: *Chanoines, Romme, Enfant rosti, Narbonne, Eaue et vergier*, and *Buef*.

The adjective *bon* also occurs six times in the *J-corpus* in the position immediately before *exemple* (16.22%). Comparing this again to our reference corpora, we find comparable rates with no statistically significant difference for the *A-corpus* (8.33%, one instance of twelve occurrences of *exemple*,  $p = 0.54603$ ) and the *P-corpus* (17.95%, fourteen of seventy-eight,  $p = 0.83887$ ) but a significantly lower rate of 4.56% for the *B-corpus* (thirty-six of 789,  $p = 0.00456$ ). This is notable in that it suggests that Jehan's use of this phrase was not particularly unusual amongst other poems written in alexandrine monorhyme quatrains, or indeed the diverse texts which were chosen for inclusion in fr. 24432, but that the use of the phrase in these three corpora 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 sixteen instances of the use of one or more of *oïr*, *raconter*, *escouter*, *dire*, *recorder*, and *(r)amenter* within the immediate context of *exemple*, resulting in a use rate of 43.24%. This time, not a single such example appears in the *A-corpus* ( $p = 0$ ), and much lower rates are observed for both the *P-corpus* (17.95%, fourteen of seventy-eight,  $p = 0.00389$ ), and the *B-corpus* (6.46%, fifty-one of 789,  $p = 0$ ). This once again reinforces the coherence of the *J-corpus*, and Jehan's apparent interest in presenting the oral nature of his storytelling. It also suggests that his conception of *exemple* was as a spoken genre, whereas this idea was not necessarily shared by the authors of the texts in the *A-corpus*.<sup>21</sup>

The didactic associations of *exemple* are also made clear by its frequent collocation with verbs which also have such connotations, including *montrer* (seven occurrences, 18.92%), *prendre* (six, 16.22%), *amender* (three, 8.12%), and *donner* (two, 5.41%). This didactic interpretation of the term does appear to be more widely shared than the association with orality, both in other narrative poems in alexandrine monorhyme quatrains in the *A-corpus*, and more widely. Half of the twelve uses of *exemple* in the *A-corpus* use such terms, with two examples each of *(de)montrer*, *prendre*, and *donner* (16.67%,  $p = 0.90122$ , 0.97818, and 0.20875 respectively). In the *P-corpus*, meanwhile, we find nine instances of *montrer* alongside *exemple* (11.54%,  $p = 0.28559$ ), eleven of *prendre* (14.10%,  $p = 0.76007$ ), none of *amender* ( $p = 0.01990$ ), and four 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 Jehan's

<sup>21</sup> For more on this, see the discussion of performance on pp. 61–62.

texts, which helps us to better understand the ways in which *exemple* and *exemplaire* were understood within his immediate literary context, as terms inextricably linked with didacticism, thus making Jehan's association of the terms with orality all the more notable.

A similar story appears in the *B-corpus*, where *montrer* appears alongside *exemple* at a rate of only 4.56% (thirty-six of 789), *prendre* at 12.17% (ninety-six), *amender* at only 0.25% (two), and *donner* at 5.70% (forty-five). 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 across 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*-, *A*-, and *P-corpora* and those written by others in different *milieux*.

All in all, we may therefore say with some certainty that our statistical consideration of *exemple* has given several 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.

## **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*. From its earliest appearances as a generic descriptor at the end of the twelfth century, the term remains remarkably difficult to define until it is reinvented and stabilized shortly before 1340, a few decades after Jehan's writing, at the hands of Guillaume de Machaut and others, when it comes to refer to "narrative verse texts in which the author's first-person literary persona recounts adventures in love" (Swift 2021, 264).<sup>22</sup>

The most comprehensive attempt at a description of the pre-1340 *dit* is that by Monique Léonard (1996), in which she critically examines the use of the term, concluding that it is not a defined generic category like the

<sup>22</sup> In addition to Swift (2021), on the post-1330s *dit* see also Cerquiglini-Toulet (2001). For a broader examination of the *dit* from the thirteenth to the fifteenth centuries, see the volume edited by Delage-Béland and Salamon (2022), and in particular the chapter by Gabriel Cholette.

*chanson de geste, fabliau, or lai*, but instead is a broad term that can be used to refer to any “pièce rimée, non chantée, plutôt brève, souvent détentrice d’un enseignement ou d’un message à transmettre au lecteur” (1996, 350).<sup>23</sup>

The term can be used of texts even if they already meet the requirements for inclusion in other categories like *miracles* and *exemples*, as well as of texts that do not sit comfortably within any other generic category, in which cases *dit* functions as what Helen Swift (2021, 265) describes as a “for want of anything better” category. Through her comprehensive numerical examination of texts which are considered *dits*, Léonard also reveals that poems can be considered *dits* regardless of whether they are strophic or non-strophic, and regardless of their metre or rhyme scheme. Indeed, she even identifies five texts, described plausibly as *dits*, which are nonetheless written in prose.

There is therefore nothing particular to the term *dit* which makes it an especially good generic descriptor for the texts of the *J-corpus*, since as we have seen their content is generally suitable for consideration as *miracles*, while Jehan appears particularly fond of the term *exemple*. And yet, it is *dit* 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, for instance, with Munk Olsen even going so far as to include the term in the title of his edition, and more recent scholarship is unanimous in its uncritical acceptance of the term to refer to the texts of Jehan’s *œuvre*.<sup>24</sup> It is therefore surprising to find that the generic self-descriptor *dit* is actually used relatively sparsely in the *J-corpus*, appearing on only sixteen occasions (Table 14), which makes it slightly less frequent than *exemple*.

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-* and *A-corpora* it has been necessary to manually determine which uses of the word form *dit* can be considered generic descriptors, and which cannot.<sup>25</sup>

---

**23** “rhymed piece which is not sung, rather brief, and often conveys a lesson or message to the reader.”

**24** See, for example, Cholette (2022).

**25** For reference, only five of the sixteen manually identified uses of *dit* as a noun in Jehan’s texts have been correctly identified by RNNTagger, so we may hypothesize, assuming the rates of failure are the same, that there are some seventy such examples to be found across the *P-corpus*, and perhaps some 1,632 examples in the *B-corpus*. For the *A-corpus*, RNNTagger is slightly more successful, at seven out of fourteen uses of *dit* as a generic descriptor tagged correctly.

Table 14. Uses of *dit* as a generic descriptor within the *J-corpus*.

| Text                        | Line | LH                                          | KW                | RH                                                     |
|-----------------------------|------|---------------------------------------------|-------------------|--------------------------------------------------------|
| <i>Respon</i>               | 4    | Weil <b>recorder</b> un                     | dit               | dont la matiere est <b>belle</b>                       |
| <i>Pommes</i>               | 333  | Par essemble vous<br>ai en ce               | dit               | <b>raconté</b>                                         |
| <i>Romme</i>                | 2    | Veil <b>dire</b> un <b>biau</b>             | dit               | qui est sanz vilenie                                   |
| <i>Chevalier et escuier</i> | 3    | Que quant je <b>dis</b> un                  | dit               | a tous generaument                                     |
| <i>Narbonne</i>             | 4    | Or vous <b>dirai</b> bon                    | dit               | qui ne me fera <b>taire</b>                            |
| <i>Cordouanier</i>          | 8    | A tiex gens n'ai je<br>cure de <b>biaus</b> | diz               | <b>recorder</b>                                        |
| <i>Enfant sauva mere</i>    | 27   | <b>Entendés</b> un <b>biau</b>              | dit               | que je vous veil <b>noncier</b>                        |
| <i>Anelés</i>               | 333  | Ainssi com vous <b>orrés</b><br>se tout mon | dit               | <b>recort</b>                                          |
| <i>Buef</i>                 | 3    | Veul <b>recorder</b> un                     | dit               | a touz communement                                     |
| <i>Buef</i>                 | 33   | Et pour ce que mon                          | dit <sup>26</sup> | soit plus ferme et estable                             |
| <i>Buef</i>                 | 36   | Qu'a vos ames puist<br>estre ce <b>biau</b> | dit               | proufitable                                            |
| <i>Buef</i>                 | 381  | Or commence <b>biau</b>                     | dit               | et plain de grant sustance                             |
| <i>Buef</i>                 | 789  | Vous qui avez <b>oÿ</b><br>ce <b>biau</b>   | dit               | <b>recorder</b>                                        |
| <i>Abesse</i>               | 1    | Ceus qui voudront oïr<br>un <b>biau</b>     | dit               | <b>recorder</b>                                        |
| <i>Abesse</i>               | 6    | Un                                          | dit               | qu'ai en propos de cest<br>miracle faire <sup>27</sup> |
| <i>Sauveur</i>              | 2    | Je vous veul<br><b>recorder</b> ung         | dit               | de grant pitié                                         |

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, alongside the sixteen nominal uses identified for the *J-corpus* in Table 14 and the fourteen uses in the *A-corpus* given in Table 15, I consider

**26** 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.

**27** Note here Jehan’s division between *dit* as a generic marker, and *miracle* as a description of an event, as observed also by Cerquiglini-Toulet (2022, 129n1).

only the twenty-two 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 therefore be cautious about drawing any firm conclusions from these latter two corpora, 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.

Table 15. Uses of *dit* as a generic descriptor within the *A-corpus*.

| Text                     | Line | LH                                          | KW  | RH                                       |
|--------------------------|------|---------------------------------------------|-----|------------------------------------------|
| <i>Cristofle</i>         | 17   | Cellui qui fist le                          | dit | a <b>recordé</b> souvent                 |
| <i>Dieudonnée</i>        | 3    | Voel <b>recorder</b> un                     | dit | de boin entendement                      |
| <i>Dieudonnée</i>        | 54   | Que qui ora mon                             | dit | jusques a le parclose                    |
| <i>Dieudonnée</i>        | 57   | La dame dont le                             | dit | vous voel ci endroit <b>dire</b>         |
| <i>Dieudonnée</i>        | 492  | Or commenche dou                            | dit | la plus piteuse istoire                  |
| <i>Grégoire</i>          | 7    | Vous diray ung <b>beau</b>                  | dit | maintenant sans targer                   |
| <i>Grégoire</i>          | 9    | Le                                          | dit | que je <b>diray</b> est de vraye matiere |
| <i>Jehan Ewangeliste</i> | 78   | S’il <b>oient</b> si bons<br>moz et si bons | diz | repandre                                 |
| <i>Jehan Ewangeliste</i> | 82   | Mes                                         | diz | et ma matere de quoi<br>j’iere rimant    |
| <i>Guillaume</i>         | 1    | Pour recorder un                            | dit | sui ci endroit venus                     |
| <i>Guillaume</i>         | 627  | Or commance du                              | dit | la piteuse matere                        |
| <i>Robert</i>            | 2    | Veul commencer un                           | dit | mais por la grant misere                 |
| <i>Robert</i>            | 479  | Tant chevaucha Robert<br>pour abregier mon  | dit |                                          |
| <i>Dieu d’Amours</i>     | 17   | Qui a <b>biaus</b>                          | diz | <b>trover</b> veut son corage<br>enbatre |

Looking at Table 15, we note first that there are two uses of *dit* as a generic descriptor in both *Guillaume* and *Robert*, the two texts which we argue are unambiguously the works of Jehan. Turning to the rest of the *A-corpus*, we observe that the particularly high frequency of *bel* immediately preceding *dit* in the *J-corpus* (seven instances of sixteen, 43.75%) contrasts with a notably but not statistically significant lower rate in the *A-corpus* (two of fourteen, 14.29%,  $p = 0.09354$ ), and statistically significant lower rates within the *P*- (two of twenty-two, 9.10%,  $p = 0.01553$ ) and *B-corpora* (seven of 510, 1.37%,

$p = 0$ ).<sup>28</sup> More strikingly, in twelve of the sixteen lines containing 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 an appearance rate of 75%. Again, the lower rate for the *A-corpus* is not statistically significant (six of fourteen, 42.86%,  $p = 0.08056$ ), but is close enough to our significance threshold as to be notable, and a statistically significant distinction is observed for the instances in the *P-corpus* in which *dit* is tagged as a noun (three of twenty-two, 13.64%,  $p = 0.00010$ ), and likewise in the *B-corpus* (seventy-seven of 510, 15.10%,  $p = 0$ ).

Thus, while Jehan does not use *dit* as often as we might anticipate based on the frequency of the term in modern scholarship to describe his texts, instead making greater use of *exemple* and its variants, the same trends appear in the words that appear alongside *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 many of the rates observed for our reference corpora, reinforces the cohesion of the *J-corpus* and strengthens our shared-author hypothesis. It also suggests that, for Jehan, the words used to describe his texts were largely interchangeable, particularly *exemple* and *dit*, and that he may have decided which word to use primarily based on the metrical requirements of its immediate context.

## 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 fr. 24432, in which most of his attributed texts survive, were less flexible in their approach. Of the twenty-two texts from within the *J-corpus* which are found in fr. 24432, all but *Pommes* and *Anelés* contain an initial *incipit* marker, and in all twenty cases, the rubricator describes the text as a *dit*. While the same term also appears in the *explicits*, it is less frequent, appearing in nine of the twenty-one *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

<sup>28</sup> Across her wider corpus of *dits*, Léonard (1996, 209) notes the appearance of *bel* with *dit* in some fifty-seven texts. She notes that texts containing this feature are often longer, but such a correlation does not appear to bear out for the *J-* or *A-corpora*.

by the *explicit*, meaning that at some point in the paratexts of fr. 24432, all twenty-two 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.<sup>29</sup> The same is true for *Robert* and *Guillaume*, which are both described as *dits* in their *explicits*, and the former in its *incipit*.<sup>30</sup> 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.<sup>31</sup> 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.

Even though 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, we ought to be careful not to read too much into this use of descriptors in fr. 24432. The term *dit* is by far the most frequent across the paratexts of the codex, being used in the *incipits* of a further thirty-two texts which are not attributed to Jehan, and appearing in the *explicits* of a further twenty-three. Allowing for overlapping usage, we therefore find that, including Jehan's attributed poems, of the eighty-five texts across fr. 24432 for which either an *incipit* or *explicit* is included, fifty-six (65.88%) are at some point described as *dits* in their accompanying paratexts.

The term *dit* thus appears to have been in far more frequent use by the compiler(s) of fr. 24432 than by Jehan himself, and it is therefore clear that

---

**29** In the paratext surrounding the two texts attributed to Jehan outside of fr. 24432, *Sauveur* (BnF, MS Ars. 2115) and *Abesse* (BnF, MS fr. 12483), no such descriptor is used, suggesting this feature may be particular to fr. 24432.

**30** The lack of an *incipit* for *Guillaume* likely stems from it being the first item in fr. 24432.

**31** Given the relative paucity of self-description using the term *dit*, we may therefore question the validity of the approach taken by Cholette (2022, 93), who selects his corpus of thirty-four “*dits non allégoriques où le caractère narratif domine*” (“non-allegorical *dits* where the narrative character is dominant”)—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. Based on our analysis, many of Jehan's poems should in fact have been excluded from Cholette's corpus.

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, but from engagement in a chain of reception initiated or followed by the compiler(s) of fr. 24432. That is not to say that such a decision is invalid, and I 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 seen across our examination of generic descriptors that Jehan has his own distinct method of describing his texts, which differs from the expectations of previous scholarship, from comparable and divergent texts, and from the perspective of the compiler(s) of fr. 24432. This distinctiveness, in and of itself, reinforces the shared-author hypothesis. Further, the fact that the compiler(s) of fr. 24432 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.

## Building the Case for the Shared-Author Hypothesis

Modern generic labels and manuscript paratextual markers can obscure authorial intent, but the internal self-descriptors found within the *J-corpus* offer compelling evidence for collective attribution. For instance, Jehan's significant preference for the phrase (*un moult tres*) *bel exemple*, alongside his association of generic terms with verbs of orality, strengthens the shared-author hypothesis. Meanwhile, *Richard's* use of the *bel miracle* structure further supports its status as a surviving echo of an earlier work by Jehan, since this feature is notably rarer in our reference corpora. Finally, while the ubiquitous use of *dit* to describe Jehan's texts does not stem from the use of the term in the texts themselves, the fact that the rubricator(s) of fr. 24432 consistently applied this term to the texts of the *J-corpus*, as well as *Guillaume* and *Robert*, suggests that early readers recognized the poems' collective similarities.

## Methodological Contribution

The approach presented in this chapter shifts from a focus on prescriptive modern labels to a *genre poetics* akin to that of Gingras (2016), which is rooted in data observed within the texts themselves, and their manuscript contexts. By using digital textual analysis to identify the distribution and collocations of specific generic descriptors, I have shown that authorial self-conceptions of genre can offer strong evidence in favour of collective attribution. Such an approach relies on humanistic engagement with a given text to situate generic self-descriptors within a given historical context. Comparing the self-descriptors in our target corpus to the uses in wider reference corpora like the *B-corpus*, I have also shown how it is possible to discriminate between broad linguistic trends and the idiosyncratic preferences of an individual author.

## TERMS OF ADDRESS

WE HAVE SEEN alongside our examination of the use of generic descriptors that the poems attributed to Jehan are 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 ("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 ("s'il vous plaisoit a taire," *Mescreant*, v. 18), and the use of terms to directly address an audience ("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,<sup>1</sup> 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,

---

I Cf. Marnette (1998), amongst others.

---

**ABSTRACT** This chapter explores how the texts we have attributed to Jehan navigate the relationship between narrator and audience, identifying idiosyncratic uses of terms of address which support the shared-author hypothesis. Through comparison with the *A-* and *F-corpora*, for example, I show that Jehan's preference for addressing his audience as *bone gent* is unusual, and provides the clearest evidence so far for the inclusion of *Robert*, *Richard*, and *Jehan Paulus* in his *œuvre*. Additionally, using digital methods I challenge previous assertions about Jehan's use of *seigneur*, showing that, when used to address an audience, it has an egalitarian sense that broadens Jehan's potential audience, rather than serving as a marker of social hierarchy. A further, particularly significant piece of evidence in support of the shared-author hypothesis appears in Jehan's formulaic use of *tous et toutes* and *ceulz et celles*, which emphasizes Jehan's aim to address all sections of society, regardless of sex or status, and which is especially unusual across our reference corpora. In this chapter, I largely exclude the *B-*, *P-*, and *R-corpora* from our analysis, since their diverse contents do not consistently address their audience in the same way as is observed for the *J-*, *A-*, and *F-corpora*.

---

un carrefour, le dimanche ou aux jours de foire" (1939, 343).<sup>2</sup> 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."<sup>3</sup>

Whether we go so far as to take all these 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 most 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.

### **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, 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,"<sup>4</sup> and enhances the associations between the texts of the *J-corpus*, and thus our hypothesis that they share an author.

Beginning first with the phrase *bone gent*, which Morawski (1939, 344) identifies as being used by Jehan to refer to his audience, we observe six instances of the phrase in this use, as seen in Table 16, 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.

Within the *A-corpus* of narrative poems written in alexandrine monorhyme quatrains which have not previously been attributed to Jehan, we observe a further twenty-eight uses of *bone gent*, of which nine are used in a vocative sense, as a term of address for an audience, as displayed in Table 17.

---

2 "undoubtedly intended for public recitation, in a square, at a crossroads, on Sundays, or on fair days."

3 "recited outdoors in a public square after mass, or on a fair day."

4 "recited outdoors in a public square."

Table 16. Uses of *bone gent* as a means of referring to the audience within the *J-corpus*.

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

Table 17. Uses of *bone gent* as a means of referring to the audience within the *A-corpus*.

| Text                | Line | LH           | KW          | RH                                      |
|---------------------|------|--------------|-------------|-----------------------------------------|
| <i>Alexis</i>       | 473  |              | Bonnes gens | mout fessoit ce saint corps a prizier   |
| <i>Jehan Paulus</i> | 61   |              | Bonnes gens | escoutés com le prudons fut pris        |
| <i>Sebastien</i>    | 214  | Et pour tant | bonnes gens | que le chevecier doit                   |
| <i>Robert</i>       | 6    |              | Bonne gent  | il n'est roy prince ne connestable      |
| <i>Robert</i>       | 313  |              | Bonne gent  | a cel temps que Robert fu el bois       |
| <i>Robert</i>       | 404  |              | Bonne gent  | qui fait mal droiz est qu'il le compere |
| <i>Robert</i>       | 664  | Par amours   | bonne gent  | oez que Robert fist                     |
| <i>Richard</i>      | 1    |              | Bonnes gens | qui avez ouÿ de maint histoire          |
| <i>Richard</i>      | 248  |              | Bonnes gens | en ce temps que vous m'oés noncer       |

In this case, we note that all but two of them appear within *Robert*, *Richard*, and *Jehan Paulus*, three of the four texts which we have proposed for inclusion in Jehan's *œuvre*. Given the frequent appearances of *bone gent* in these texts now attributed to Jehan, it is striking that we do not observe the same frequency of use in the rest of the *A-corpus*, nor indeed in the *F-corpus* of *fabliaux*, which also contain numerous direct appeals to their audiences, including requests to come close and for silence. Across the 101 texts of the latter corpus, with a total word count greater than that of either the *J-* or *A-corpus*, there are only six uses of *bone gent*, of which none are used

to refer to the text's audience.<sup>5</sup> This suggests that the term was not a particularly standard means of addressing an audience in a text intended for performance, and that the texts in which we have observed this feature are unusual in their vocative use of *bone gent*.

We may therefore argue that Jehan had a clear, individual preference for this phrase when used to address an audience, resulting in arguably the clearest evidence in favour of the shared author hypothesis so far, and adding to our growing body of evidence for the inclusion of *Robert*, *Richard*, and *Jehan Paulus* in Jehan's *œuvre*. While these indicators alone are not enough to reach certainty on the attribution of these texts to Jehan, they certainly point in that direction, and in our later analysis we will elucidate further areas where the texts display features which overlap with the pre-attributed texts of the *J-corpus*, supporting a shared attribution.

To return for now, however, to the *J-corpus* of poems previously attributed to Jehan, it is worth also considering the uses of *bone gent* beyond references to a text's audience. By examining the contexts of these twelve such instances, presented in Table 18, we can 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-* and *A-corpora* 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 16 or Table 17.

Jehan's uses of *bone gent* to refer directly to his audience therefore differ from more narrative ones, but the latter still have some utility in helping us comprehend the apparent possible meanings of this phrase for Jehan. It is possible to distinguish sub-groups within the narrative uses, again presented in Table 18. In the first category (A), we find four examples which

---

**5** Beyond the *J-* and *A-corpora*, notable use of *bone gent* appears within *Le Dit de Menage*, another poem in alexandrine monorhyme quatrains found within fr. 24432, which contains four instances of the phrase, of which three refer to the text's audience. For an edition and brief study, see Dows-Miller (2025a). The poem 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 uses of *bone gent*, while evidence of a preference for the author of *Le Dit de Menage*, are not enough to support shared authorship with the *J-corpus*.

Table 18. Uses of *bone gent* within the *J-corpus* for purposes other than to address the audience.

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

represent an oblique method of referring to the audience using the third-person, and are thus comparable to the examples in Table 16 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 three out of four cases. It is notable also that the uses in category A all appear in prologues, and in three 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 forms which refer in a general way to those who are morally “good,” yet suffer from 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 universalizing 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*),<sup>6</sup> 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.

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 overall the term is used in a broader sense to apply to his whole audience. The consistency of the use of *bone gent* across many of the texts of the *J-corpus* therefore bind the poems together in a way that strengthens the sense that they are the work of a single author. Further, the frequent use of what appears to be an egalitarian term 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 sections of society.

Comparing the *J-corpus* to the nineteen narrative uses of *bone gent* in the *A-corpus*, we find that category A uses appear to be a particular feature of Jehan's writing, since no such structure appears in the texts of the *A-corpus*. The inverse is true, meanwhile, for instances in which *bone gent* is used vocatively by a character within the narrative to address an assembled audience, which does not appear in the *J-corpus* but appears on five occasions in the *A-corpus* (once in *Dieudonnée*, and four times in *Jehan Ewangeliste*). There are just two category B uses in the *A-corpus*, with the remaining twelve examples falling within category C, both of the high- and low-status variety.

In keeping with the vocative uses of *bone gent* we observed in *Robert* in Table 17, the narrative uses of this phrase also align with those of the *J-corpus*, reinforcing the sense that *Robert* is also the work of Jehan. Within the text, we find 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. We once again get the impression that the author behind *Robert* had a similar interpretation of *bone gent* to Jehan, reinforcing the suggestion that

---

<sup>6</sup> 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."

these are indeed the same person. A category C narrative use of the phrase also appears in *Guillaume*, the other poem of the *A-corpus* that appears in fr. 24432, and which we also now attribute to the same author, meaning that examples of *bone gent* appear in uses aligned with the *J-corpus* across all four of our newly attributed texts.

### **seigneur**

Alongside Morawski's observation of the use of *bone gent*, Munk Olsen (1978, xiii) also argues that Jehan's texts are characterized by the presence of several 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).<sup>7</sup>

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 universalizing way, and it does also differ somewhat from Nykrog's argument that *seigneur* was often used in the *fabliaux* as a general term, referring "indifféremment à un groupe de vilains, de bourgeois ou de nobles" (1957, 21).<sup>8</sup> As we shall see from a digitally facilitated 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 of society.

Looking at Table 19, we find that Jehan's uses of *seigneur* are notable in their rarity, despite Munk Olsen's insistence on their significance to the *J-corpus*.<sup>9</sup>

<sup>7</sup> "to designate an audience of peasants or bourgeois while flattering them a little."

<sup>8</sup> "indiscriminately to a group of peasants, bourgeois, or nobles."

<sup>9</sup> 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 "[i]l 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'" ("it should not be forgotten that the *Dits* contain almost 5,000 lines, so a word that appears two or three times is not necessarily a word 'that the author uses a lot,'" 1963b, 155n17).

There are ninety-one uses of the word across Munk Olsen’s corpus of Jehan’s poems, yet only three are direct appeals to the audience.<sup>10</sup>

Table 19. Uses of *seigneur* as a means of referring to the audience within the *J-corpus*.

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

The remaining eighty-eight instances occur within the tales themselves, either during moments of narration (sixty-one, 70.45%) or direct discourse (twenty-seven, 29.55%). Of the examples used in direct discourse, twenty-two 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.<sup>11</sup> We may separate these examples into two categories of usage: hierarchical and egalitarian. Instances of the former category, found in Table 20, 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.

Table 20. Hierarchical uses of *seigneur* within the direct discourse of the *J-corpus*.

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

**10** It is worthy of note that *Robert* matches this usage, containing ten uses of *seigneur* of which none appear in a direct address to the audience. *Guillaume* contains a further twenty, which again do not appear in a vocative sense.

**11** Of the five uses of *seigneur* within direct discourse which are not used as terms of address, there are two references to God (*Juitel* and *Mescreant*), two to a woman’s husband (*Florence* and *Anelés*), and one to a speaker’s *suzerain* (*Pommes*).

In the category of egalitarian uses of *seigneur* in Table 21, 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 21. Egalitarian uses of *seigneur* within the direct discourse of the *J-corpus*.

| Text                            | Line | LH                             | KW        | RH                                       |
|---------------------------------|------|--------------------------------|-----------|------------------------------------------|
| <i>Respon</i>                   | 150  | Et dient l'un<br>a l'autre "   | Seigneurs | alons nous en                            |
| <i>Respon</i>                   | 171  | "[...]                         | Seigneurs | je sui vos sire par Dieu de paradis"     |
| <i>Respon</i>                   | 176  | "                              | Seigneurs | " dist le chanoinne le roi de maesté     |
| <i>Chanoignes</i>               | 33   | "                              | Seigneurs | par celle foi que vous a Dieu devez      |
| <i>Chanoignes</i>               | 47   | "                              | Seigneur  | " dist li ainsnés "entendez mon recort   |
| <i>Chevalier<br/>et escuier</i> | 50   | "                              | Seignor   | comment vous est" ce dit li ennemis      |
| <i>Florence</i>                 | 103  | Quant son frere<br>leur dist " | Seignours | lessiez ester"                           |
| <i>Florence</i>                 | 122  | "                              | Seignours | " dist il "pour Dieu la vengeance prenez |
| <i>Anelés</i>                   | 221  | "Biau                          | seigneurs | vous veés ci endroit en present          |
| <i>Anelés</i>                   | 273  | Et dist "                      | Seigneurs | je jure par les sains qui sont ci        |
| <i>Anelés</i>                   | 309  | "                              | Seigneurs | or entendez pour Dieu granz et petiz     |
| <i>Anelés</i>                   | 410  | "                              | Seigneurs | ne savés pas pourquoi vos ai mandés      |
| <i>Anelés</i>                   | 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 (thirteen of twenty-two). Indeed, six of the nine 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, as he does in the three examples in Table 19, we may thus 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 all individuals to come and hear his moral lessons.

Table 22. Uses of *seigneur* as a means of referring to the audience within the *A-corpus*.

| Text            | Line | LH          | KW        | RH                     |
|-----------------|------|-------------|-----------|------------------------|
| <i>Basile</i>   | 1    | Or escoutés | seigneur  | que Dieus vous beneïe  |
| <i>Grégoire</i> | 1    | Or entendez | seigneurs | que Jhesus vous beneye |

As for how unique this feature is to Jehan’s texts, we find that uses of *seigneur* by the narrator to address his audience are similarly rare in the *A-corpus*, with only two such examples, as in Table 22.<sup>12</sup> Of the twelve vocative uses of the term within direct discourse, here the examples are split equally between hierarchical and non-hierarchical uses, with all five of the singular forms tending towards the former, and all but one of the seven plural forms towards the latter.

Turning also to the *F-corpus*, we find that this 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),<sup>13</sup> while also supporting our observation that in short narrative texts, *seigneur* is not

**12** We may wonder in passing, given the similarity of these two lines, whether *Basile* and *Grégoire* share an author, even if this is not Jehan. The full stanzas bear comparison in their entirety:

|                                                                                                                                                                                             |                                                                                                                                                                                            |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Or escoutés , seigneur! Que Dieus vous beneïe,<br>Le glorieus dou ciel, li fils sainte Marie!<br>Je vous voel raconter de la vierge Marie<br>Une mout haute istore qui doit bien estre oïe. | Or entendez seigneurs, que Jhesus vous beneye,<br>Le glorieux du ciel, filz de sainte Marie!<br>Si orez de Gregoire, et l'ystoire et la vie.<br>Pour paradis avoir, souffry grande hachie. |
| ( <i>Basile</i> , vv. 1–4)                                                                                                                                                                  | ( <i>Grégoire</i> , vv. 1–4)                                                                                                                                                               |

**13** “indiscriminately to a group of peasants, bourgeois, or nobles.”

particularly widespread as a term of address for one's audience.<sup>14</sup> Of the 174 instances that appear across the *F-corpus*, only sixteen are used directly by a narrator. As with the *J-* and *A-corpora*, while the examples within direct speech in the *F-corpus* to address an interlocutor vary from distinctly hierarchical uses, such as addresses to God, to sardonically polite non-hierarchical uses against a group of ne'er-do-wells, the non-hierarchical uses are predominantly in the plural. We thus gain a similar interpretation to that found in the *J-* and *A-corpora*, 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 original 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.

### **tous et toutes and ceulz et celles**

This universality of Jehan's intended audience, inclusive of all who wish to listen, is emphasized 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 23, which reinforce that his moral messages are for the betterment of both sexes, and therefore of society at large.<sup>15</sup>

The position of all but one of these phrases in the final three syllables of the first hemistich, often following the phrase *si/or prions*,<sup>16</sup> 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 *tous et toutes* is that no examples of the same phrase appear in the *F-* or *R-corpora*, and while there are some thirty-eight examples in the

<sup>14</sup> Anecdotally, at least, it does appear to be more common in the *Chansons de geste*.

<sup>15</sup> 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 in *Buef* (v. 781), as well as the comparable phrase "chascuns et chascune" in *Chanoines* (v. 2), *Povre chevalier* (v. 159), and *Eaue et vergier* (v. 162). Examples of such universalizing phrases are thus found in eleven of the twenty-four poems of the *J-corpus*.

<sup>16</sup> Morawski (1939, 330) also observes as particularly frequent use of this phrase in Jehan's epilogues. Indeed, some kind of first-person singular imperative (usually *prions*) appears in twenty-one of the twenty-four texts of the *J-corpus*. The texts in which it does not appear are *Narbonne*, *Enfant sauva mere*, and *Merlin*.

*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.<sup>17</sup> Jehan therefore appears to be particularly unusual in his use of these terms to refer to his audience.

Table 23. Uses of *tous et toutes* and *ceulz and celles* as means of addressing the audience in the *J-corpus*.

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

Turning to the *A-corpus*, we observe eleven such structures, of which six are used to directly address the audience, as in Table 24. Of these, it is once again notable that examples are found within *Guillaume*, *Jehan Paulus*, and *Richard*, further reinforcing the attribution of these texts to Jehan.

Table 24. Uses of *tous et toutes* and *ceulz and celles* as means of addressing the audience in the *A-corpus*.

| Text                | Line | LH                    | KW              | RH                   |
|---------------------|------|-----------------------|-----------------|----------------------|
| <i>Barbe</i>        | 446  | Entendés              | tous et toutes  | ne vous retardés mye |
| <i>Jehan Paulus</i> | 367  | <b>Si prions</b>      | touz et toutes  | a la Vierge Marie    |
| <i>Confesseur</i>   | 141  | Pour Dieu escoustez   | tous et toutes  | ma raison            |
| <i>Guillaume</i>    | 2    | Dieu gart touz        | ceulz et celles | dont seray entendus  |
| <i>Richard</i>      | 816  | <b>Or prions</b>      | tous et toutes  | au Roy de Paradis    |
| <i>Déz</i>          | 120  | Quar i nous convendra | tous et toutes  | mourir               |

*Guillaume*, the first text in fr. 24432, contains an instance of *ceulz et celles* which is especially comparable to the similar structure found in *Florence*, another text of the “long-text” sub-group, both in its content invoking God’s protection for the speaker’s audience, and also in the fact that the structure

<sup>17</sup> From the late-fifteenth-century *Ressource de la chrestienté*.

appears at the beginning of the poem. This use of the phrase, while again not enough to confirm authorship on its own, therefore adds to the evidence for the claim that *Guillaume* ought to sit within Jehan's *œuvre*.

Looking beyond *Guillaume*, the most striking of the examples in Table 24 are the two uses of *tous et toutes* in *Jehan Paulus* and *Richard*, which align almost exactly with examples observed in *Chanoines*, *Cordouanier*, *Juitel*, and *Mescreant* in the *J-corpus*. In all six examples across both corpora, the structure appears in the final stanza of the poem, and in the case of *Jehan Paulus* and *Cordouanier*, forms part of a larger parallel structure, first observed by Morawski (1940, 515–16):

Si prions, touz et toutes, a la Vierge Marie,  
Qui en ces flans porta le disgne fruit de vie,  
Que tel oroyson soit de nous bouches parties,  
Qu'avoir puissions des anges la sainte compaygnie.

(*Jehan Paulus*, vv. 367–70)

Si prions, touz et toutes, a la Vierge Marie,  
Qui en ces flans porta le tres dous fruit de vie,  
Qu'il nous doint si ouvrier en ceste mortel vie,  
Par quoi aions des angres la sainte compaignie.

(*Chanoines*, vv. 409–12)

A similar parallelism is observed between the exhortatory epilogue of *Richard*, containing the same phrase, and that of *Pecherresse* within the *J-corpus*. While the epilogue to *Pecherresse* does not make use of *tous et toutes*, and the points of overlap here are more echoes than outright copies, the final lines of both poems do still share notable similarities:

[...]  
Or prions tous et toutes au Roy de Paradis,  
Et a la douce Vierge, qui est de moult hault pris,  
Que tous puissions avoir des saintz cieulx les delis.

(*Richard*, vv. 816–18)

Prions a cele Dame, qui tant est de hault pris,  
Qu'ele soit nostre escu entre nos ennemis,  
Et nous doinst si ouvrier et en fais et en dis,  
Que nous puissions conquerre nos lieux en paradis.

(*Pecherresse*, vv. 149–52)

Morawski (1940) identifies both points of similarity, but excludes *Jehan Paulus* and *Richard* from Jehan's *œuvre* based on their use of possessive adjectives. We will discuss this linguistic question in the following chap-

ter below, but for now suffice it to say that it is indeed a cogent argument against the inclusion of these texts in Jehan's *oeuvre* in their current form. Of course, we ought also to note in these examples that it is plausible that the similarities are due to copying and influence between authors, rather than one author reusing previous material (in the case of *Jehan Paulus* and *Chanoines*) or a case of self-influence (as in *Richard* and *Pecherresse*).

The gradual accumulation of evidence that will proceed throughout the remainder of this study, however, which we have already encountered in part, reveals that *Jehan Paulus* and *Richard* share a surprising number of other similarities with the *J-corpus*. We must, therefore, consider a third way between Morawski's outright rejection of these texts and a straightforward single-author hypothesis that would be suggested by these data alone: *Richard* and *Jehan Paulus*, as they survive today, cannot simply be the work of Jehan, rather both are lightly reworked versions of earlier texts which Jehan did compose, produced by some later poet. They are thus echoes of other works, alerting us to a once more expansive poetic repertoire which has since been lost. It is this hypothesis, as we shall observe, that enables us to account for why the poems align with the *J-corpus* in striking ways on some metrics, and yet are more divergent in others. As we do so, it will be possible to pass broader comment on what elements of Jehan's earlier texts were considered suitable for *remaniement*, and which were not.

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 chapters, an important observation for our overarching question of the attribution of authorship remains the unity brought to the *J-corpus* by Jehan's 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 chapter has also greatly increased the volume of evidence in favour of direct inclusion of *Guillaume* and *Robert* in Jehan's *oeuvre*, alongside the status as later versions of texts composed by Jehan granted to *Richard* and *Jehan Paulus*. The evidence given so far is not enough on its own to prove the shared-author hypothesis,

but even by identifying these texts as possible candidates for inclusion, we have also shown the value of the 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.

## Building the Case for the Shared-Author Hypothesis

The analysis of terms of address presented in this chapter provides the most significant evidence for the shared-author hypothesis thus far. The *J-corpus* displays idiosyncratic use of *bone gent* as a term by which to address an audience, and this is also replicated in *Robert*, *Richard*, and *Jehan Paulus*, adding further arguments in support of both the cohesion of the *J-corpus*, and the inclusion of these additional texts within Jehan's *œuvre*. This evidence is compounded by the formulaic use of *tous et toutes* and *ceulz et celles* in the *J-corpus* as well as in *Guillaume*, *Jehan Paulus*, and *Richard*, which is a phrase that Jehan appears to have especially favoured, but which is largely absent elsewhere. The use of these structures, meanwhile, brings our attention back to the near-identical closing stanzas shared between *Jehan Paulus* and *Chanoines*, as well as notable parallels between *Richard* and *Pecherresse*. These examples reveal the presence of Jehan's authorial voice within these previously unattributed texts even if, as appears likely for *Richard*, the surrounding material has been subsequently reworked.

## Methodological Contribution

The analysis of this chapter moves beyond a purely qualitative examination of orality, quantifying the frequency of use of terms of address, and testing them against reference corpora containing with similar oral features. My analysis demonstrates that terms of address are not simply interchangeable filler phrases, but can instead be highly specific indicators of an author's relationship with their audience. Further, by combining a data-driven approach with the close reading of individual examples, I have shown how the examination of vocative terms used within a narrative can inform our understanding of the function of such terms when they are used to address an audience.



## POSSESSIVE ADJECTIVES

AS WE NOTED in the previous chapter, there is a cogent argument, first made by Morawski (1939, 334; 1940, 506, 509, 516, 518), that Jehan's work is united by its use of possessive adjectives. In his smaller corpus of poems which he attributes to Jehan, Morawski notes that there are no instances of the analogical use of the masculine possessive adjectives *mon*, *ton*, or *son* before a feminine singular vowel-initial noun. Instead, Jehan appears to consistently prefer to use the earlier system involving the elided forms *m'*, *t'*, and *s'*. This is despite the fact, he argues, that the analogical use of *mon* etc. was increasingly possible in several dialects at the time that Jehan wrote his texts. For Morawski, this is the most important linguistic feature of Jehan's *œuvre*, and in his eyes unifies the corpus to the exclusion of possible external contenders. 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

---

**ABSTRACT** In this chapter, I address the linguistic feature taken by Joseph Morawski (1939, 1940) to be definitive evidence for the exclusion of texts from Jehan's *œuvre*, namely his apparent non-use of *mon*, *ton*, or *son* before feminine singular vowel-initial nouns. Digital analysis reveals that Jehan is indeed consistent in his preference for elided *m'*, *t'*, and *s'*, and that this consistency is shared in *Guillaume* and *Robert*, supporting their direct attribution. However, while Morawski and Munk Olsen (1963b) view Jehan's preference for elided *ma* etc. as a notable archaism, examination of the *B1-corpus* (a subset of the *B-corpus* contemporary with the *J-corpus*) and the *P-corpus* reveals that Jehan's practice is much more in keeping with fourteenth-century linguistic norms than has previously been assumed. In the cases of *Jehan Paulus* and *Richard*, I identify the variable use of metrically consistent *mon* etc. alongside elided *ma* etc. with feminine vowel-initial nouns, but rather than assuming, as Morawski did, that they must therefore be definitively excluded from Jehan's *œuvre*, I argue instead that, in the context of all the points of overlap observed up to this point, the evidence here supports the suggestion that these two texts are echoes of poems written by Jehan which subsequently underwent significant *remaniement*.

---

of *mon*, etc. before a feminine vowel-initial noun was “tout à fait courant,”<sup>1</sup> and that the absence of this feature is therefore a peculiarity of the *J-corpus*.

Of course, finding an instance of *mon* etc. before a feminine vowel-initial noun would not be enough to prove that a text was not written by Jehan. The preferences and linguistic practices of an 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, Morawski’s observation regarding Jehan’s preference for elided *ma*, etc. does appear to be borne out by a holistic analysis of the *J-corpus*: in general, elided *ma* etc. appears before all feminine vowel-initial nouns (Table 25), and *mon* etc. is restricted to use before masculine nouns (Table 26).

Table 25. Examples of elided *ma* etc. preceding feminine vowel-initial nouns within the *J-corpus*.

| Text                   | Line | LH     | KW | RH                                    |
|------------------------|------|--------|----|---------------------------------------|
| <i>Deux chevaliers</i> | 106  | A toi  | m’ | ame et mon corps commant et a Jhesu   |
| <i>Buef</i>            | 617  | “Par   | t’ | ame pren y garde” et celle va tantost |
| <i>Mescreant</i>       | 12   | Jamais | s’ | ame n’avra vers Dieu redemption       |

Table 26. Examples of *mon* etc. preceding masculine vowel-initial nouns within the *J-corpus*.

| Text               | Line | LH                      | KW  | RH                                       |
|--------------------|------|-------------------------|-----|------------------------------------------|
| <i>Pecherresse</i> | 20   | Nient plus qu’a         | mon | enfant ne li voudrai faillir             |
| <i>Narbonne</i>    | 196  |                         | Ton | enfant te vois querre en santé et en vie |
| <i>Pommes</i>      | 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.<sup>2</sup> The only case for which this is not true is a single instance of “mon erreur” found in *Chevalier hermite* (v. 48). The 2020 *Dictionnaire du Moyen Français* (DMF) lists this

1 “completely normal”

2 The term *affaire*, 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* elsewhere in the corpus.

noun as variable between masculine and feminine, and 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 three uses of *erreur* across the *J-corpus*, given in Table 27, we can see that there are no surrounding morphosyntactic clues that could provide this information. Even the uses of “grant” in *Pommes* are not definitive since, as we saw in our discussion of the dating of Jehan’s texts, within the *J-corpus* this form is generally favoured over the analogical form *grande* which later comes to replace it.

Table 27. Uses of *erreur* within the *J-corpus*.

| Text                     | Line | LH                                 | KW     |
|--------------------------|------|------------------------------------|--------|
| <i>Pommes</i>            | 162  | Du preudomme dirai qui fu en grant | errour |
| <i>Pommes</i>            | 265  | Li varlés retorna qui fu en grant  | errour |
| <i>Chevalier hermite</i> | 48   | Veil aler a confesse regehir mon   | errour |

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. In the *J-corpus*, at least, Jehan’s preference for use of the elided *ma* etc. therefore appears to be very strong, even if occasional use of *mon* etc. before feminine vowel-initial nouns is a possibility.

However, even if we accept that Jehan is consistent in his non-use of *mon* etc. before vowel-initial feminine nouns, Morawski and Munk Olsen appear not to have been entirely accurate in their claim that, by the time the texts of the *J-corpus* were composed, around the turn of the fourteenth century, *mon* etc. and *ma* etc. were already being used interchangeably before vowel-initial feminine nouns, in a way that would mark Jehan out amongst his contemporaries.

While surprisingly little linguistic scholarship exists on this feature, Lefèvre et al. (1976–1983, 2:§220) note that as Old French moves towards Middle French, in the period in which Jehan is writing, “les féminins *ma* et *sa* s’élident encore [devant voyelle] malgré un recours de plus en plus fréquent aux masculins *mon*, *son*.”<sup>3</sup> Marchello-Nizia et al. (2020, 706, 710) offer a little more detail, suggesting that over the course of the fourteenth century “les

<sup>3</sup> “the feminine forms *ma* and *sa* still elide [before a vowel] despite increasingly frequent recourse to the masculine forms *mon* and *son*.”

formes atones élidées sont remplacées analogiquement par les formes du masculin,”<sup>4</sup> but that this process of evolution depends on the noun which the possessive is describing. For example, they suggest that the elided form of *ma* etc. is common before *ame*, *amie*, and *amour* 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. A clear example of this trend comes from *Barbe* in the *A-corpus*, itself a sixteenth-century copy of either a fourteenth- or fifteenth-century text,<sup>5</sup> in which we observe the following line:

Si dist: “Barbe, ma fille, mon espouse et m’amy,  
(*Barbe*, v. 415)

If we take the texts of the *J-corpus* as having been written at the turn of the fourteenth century, we should therefore not be surprised to see no forms of *mon*, etc. before feminine vowel-initial nouns, even those for which the possessive system progressed more quickly, since this use is not yet as “courant” (“normal”) as Munk Olsen suggests (1963b, 155). This is especially true 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 linguistic contexts, evaluating both Morawski and Munk Olsen’s claims, as well as those of the linguistic scholarship.

Turning initially to the temporally broader *B-corpus*, we find four verse texts which were composed in the first quarter of the fourteenth century, and therefore are broadly contemporary with the *J-corpus*. 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 sixty-two thousand words (the *B1-corpus*), roughly comparable in length to the *J-corpus*. The diatopic variation present in the *B1-corpus* also 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.<sup>6</sup>

<sup>4</sup> “the atonic elided forms are analogically replaced by the masculine forms.”

<sup>5</sup> The manuscript in question, the only one known to contain this text, is Avignon, Bibliothèque municipale, MS 615.

<sup>6</sup> 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

Across the *B1-corpus*, there are forty-seven instances of an elided form of *ma* etc. appearing before a feminine vowel-initial noun, and fifty-five 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. There are thus no examples of *mon* etc. before unambiguously feminine vowel-initial nouns in the *B1-corpus*. As such, these texts fit the trend observed across the *J-corpus* 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.

Further evidence in support of this comes from the *P-corpus*,<sup>7</sup> containing all the texts of fr. 24432 except for those within the *J-corpus*. Examination of this corpus reveals that of all 215 instances of *mon* etc. appearing before a vowel-initial noun, only nine of these nouns are forms which, according to the DMF, are almost certainly feminine. These examples are given in Table 28.<sup>8</sup>

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*, just as in the *J-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 had an impact on the metre of verse lines.

Indeed, it would be tempting to hypothesize that scribal influence could have operated in the other direction since, as we argued in the introduction, fr. 24432 was likely 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.

---

both Picard, while the second two texts are Burgundian and Norman respectively, giving a good spread across what is now northern France.

**7** Comparison with the *F-* or *R-corpora* would be unhelpful here, given these texts largely predate the *J-corpus*, even though on occasion copies of these texts appear in the same manuscripts.

**8** It is notable that five of these nine 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., but only in part, with the process starting later but 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.

Table 28. Examples of *mon* etc. before unambiguously feminine vowel-initial nouns in the *P-corpus*.

| Folio | Form  | LH                               | KW  | RH                                   |
|-------|-------|----------------------------------|-----|--------------------------------------|
| 32va  | Verse |                                  | Son | espie <sup>9</sup> que moult amoit   |
| 50rb  | Verse | Que il voudra semer              | son | orge                                 |
| 91ra  | Verse | [...] ne plus fine,              | Son | aime <sup>10</sup> pour biauté [...] |
| 266rb | Verse | Mais por miex                    | mon | aaise avoir                          |
| 321ra | Prose | [...] acorde son cors et         | son | ame a Dieu [...]                     |
| 372va | Prose | [...] qui fust bien de luy et    | son | acointe [...]                        |
| 373rb | Prose | [...] et de respoitier           | son | ame ausint come [...]                |
| 375rb | Prose | [...] le sauvament de            | mon | ame [...]                            |
| 379va | Prose | [...] donne son cors au monde et | son | ame au deable [...]                  |

On the contrary, in fact, none of the scribes in fr. 24432 appear to have attempted to make any consistent modifications to this feature within the texts that they copied.

Table 28 contains only six such examples from prose across the whole of fr. 24432, 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. Even those examples that do appear are seemingly not the result of the preferences of the scribes of fr. 24432. For example, four of the nine examples appear within a few folios of each other (fols. 372va–379va), despite the text containing these examples spanning some fifty-eight folios,<sup>11</sup> where numerous elided forms of *ma* etc. are also found. From the palaeographic evidence it is clear that the scribe who copied this text was a single person, and so the appearance of *mon* etc. between folios 372 and 379 suggests that while the scribe himself preferred elided *ma* etc., the comparable section of his exemplar contained greater use of *mon* etc. before vowel-initial feminine nouns. While there were therefore likely scribes more consistent in their use of *mon* etc. before vowel-initial feminine nouns whose work was circulating in the same context as that in which fr. 24432 was produced, the overwhelming preference within the manuscript does appear to be for the use of elided *ma* etc.

<sup>9</sup> *espie*, “spy.”

<sup>10</sup> As in *ame*.

<sup>11</sup> The text is a vernacular translation of the *Elucidarium* known as *Lucidaire*, edited in Türk (2000).

The scribes of fr. 24432 therefore preferred to retain or introduce elided *ma* etc. before feminine vowel-initial nouns, but this seemingly does not mean that the usage in earlier versions of Jehan's *œuvre* was more variable. Evidence from later copies of Jehan's texts found outside of fr. 24432 reinforces the suggestion that the author's preference was for elided *ma* etc. regardless of the preferences of the scribes in fr. 24432. Unlike many texts of the *J-corpus*, which are generally *unica*, *Florence* also appears within Grenoble, BM, MS 319, a composite manuscript dating to the fifteenth century. This witness is thus later than fr. 24432, and contains three examples of *mon* etc. used before feminine vowel-initial nouns, presented in Table 30.

Table 29. Uses of *mon* etc. before vowel-initial feminine nouns in the copy of *Florence* found in Grenoble, BM, MS 319, compared to fr. 24432.

| Line | Grenoble, BM, MS 319                                       | fr. 24432                                                 |
|------|------------------------------------------------------------|-----------------------------------------------------------|
| 15   | En Dieu servir avoit toute mise <b>son</b> entente         | En Dieu servir avoit bien mise s'esperance                |
| 253  | "En vostre sainte garde commant mon cors et <b>mon</b> ame | "En vostre sainte grace commant mon cors et <b>m'</b> ame |
| 392  | Et puis saicha <b>son</b> espee qui fu de bonne forge      | Puis a sachié l'espee qui fu de bonne forge               |

In each case, the elided form is evidently the preferable option, given that the line containing *mon* etc. otherwise becomes hypermetric. It is therefore clear that Jehan's preference was for elided *ma* etc. in such cases, while the scribe of Grenoble, BM, MS 319 (or one earlier in the chain of transmission) intervened in the text, updating the language to one which more accurately reflected his *scripta*, even at the expense of a coherent metre.

Beyond the case of this one manuscript, it is notable that this preference among later scribes for the modification of elided *ma* etc. to *mon* etc. does not appear with any consistency. *Sauveur* and *Abesse*, the two texts of the *J-corpus* not to appear within fr. 24432, display no use of *mon* etc. before feminine vowel-initial nouns, nor does the additional copy of *Anelés* found in the late fifteenth-century manuscript Paris, Bibliothèque nationale de France, MS français 5036. While it was therefore certainly possible for later scribes to modify Jehan's use of this feature to fit their own linguistic practices, this is not a universal trend, and in the earliest copies of the texts of the *J-corpus*, it is almost certainly true that elided *ma* etc. predominated before feminine vowel-initial nouns.

This observation proves useful when considering those texts outside the *J-corpus* to which we have also extended the shared-author hypothesis. *Guillaume* and *Robert* do not contain examples of *mon* etc. before a feminine

vowel-initial noun, reinforcing the validity of the attribution of these two texts to Jehan in their surviving form. *Jehan Paulus* and *Richard*, however, display more varied use, as presented in the tables below. Morawski is therefore right to observe, as he does for both these texts (1940, 516, 518), that they apparently display a use of possessive adjectives which is different to that of the *J-corpus*.

Table 30. Examples of *mon* etc. before unambiguously feminine vowel-initial nouns in *Jehan Paulus* and *Richard*.

| Text                | Line | LH                              | KW  | RH                                       |
|---------------------|------|---------------------------------|-----|------------------------------------------|
| <i>Jehan Paulus</i> | 328  | Emporterent                     | son | ame en chantant liement                  |
| <i>Jehan Paulus</i> | 351  | De la Vierge Marie et en        | son | oroyson                                  |
| <i>Jehan Paulus</i> | 360  | Enterrer fist l’ermite dessoubz | son | oroyson                                  |
| <i>Jehan Paulus</i> | 127  | Lors quiert                     | son | oroyson et crie hautement                |
| <i>Jehan Paulus</i> | 146  | Il quiert                       | son | oroyson mais il l’avoit gettee           |
| <i>Jehan Paulus</i> | 343  | Osta                            | son | oroyson quant il l’ot pronuncee          |
| <i>Jehan Paulus</i> | 162  |                                 | Son | oroyson regrette dolent est qu’il ne l’a |
| <i>Richard</i>      | 314  | Dedans                          | son | abbaye dont party fut tout seulx         |
| <i>Richard</i>      | 304  | Mais s’en                       | son | abbaye tout droit arriere va             |
| <i>Richard</i>      | 285  | S’en vault avoir                | son | ame mais il n’y a povoir                 |
| <i>Richard</i>      | 557  | Or avra ja Richart le diable en | son | aye                                      |

Table 31. Examples of *ma* etc. before feminine vowel-initial nouns in *Jehan Paulus* and *Richard*.

| Text                | Line | LH                                      | KW | RH                            |
|---------------------|------|-----------------------------------------|----|-------------------------------|
| <i>Jehan Paulus</i> | 144  | Que tel penance face que                | s’ | ame soit sauvee               |
| <i>Richard</i>      | 276  | Et quant il fut levé de                 | s’ | abbaye yssi                   |
| <i>Richard</i>      | 303  | S’il s’en va sur                        | s’ | amie le diable l’avra         |
| <i>Richard</i>      | 291  | Se noya tout ainsi qu’il alloit sus     | s’ | amye                          |
| <i>Richard</i>      | 681  | Et croy selon                           | m’ | entente que j’eüsse esté mors |
| <i>Richard</i>      | 474  | Je te pourfenderay de                   | m’ | espee esmoulue                |
| <i>Richard</i>      | 175  | Que je ne luy donnasse de               | m’ | espee le trenchant            |
| <i>Richard</i>      | 556  | Que nulz homs vous assaille que a       | m’ | espee n’octie                 |
| <i>Richard</i>      | 425  | Certes c’est covardise d’en avoir trait | s’ | espee                         |
| <i>Richard</i>      | 406  | Fors que tant que                       | s’ | espee sacha le duc courtoys   |
| <i>Richard</i>      | 247  | Alla sacher                             | s’ | espee tout le detrancha lors  |
| <i>Richard</i>      | 389  | Et Richart demoura pour veiller         | s’ | espousee                      |

Across both texts, there are eleven examples of *mon* etc. used before an unambiguously feminine vowel-initial noun, and twelve of the inverse, where elided *ma* etc. is used. *Jehan Paulus* displays a preference for *mon* etc. over elided *ma* etc., while the inverse is true for *Richard*. It is especially notable that in the case of *ame* in *Jehan Paulus* and *abbaye* in *Richard*, the usage fluctuates, with at least one example of each, suggesting at least the possibility that the uses of *mon* etc. before unambiguously feminine nouns are the result of later modifications to the text. Indeed, it is possible for all of the examples in Table 30 to hypothesize an earlier form involving elided *ma* etc., which a later *remanieur* could have modified either by shifting from diæresis to syneresis, removing filler words, choosing not to pronounce the *e-muet* at the caesura (which occurs elsewhere in Jehan's poetry), or performing elision elsewhere in the line, thus allowing them to make use of *mon* etc. before a feminine vowel-initial noun, which better fitted their *scripta*.

As we have noted, *Jehan Paulus* appears in the same manuscript as the later copy of *Florence* that contains *mon* etc. before feminine nouns, and so we may be tempted to suggest that scribal preference was also at play here. However, even though the modifications necessary to adjust the language of *Jehan Paulus* from a putative earlier version are not especially drastic, there is evidence in favour of a more significant reworking than small scribal modifications alone. In *Florence*, the scribal modifications resulted in hypermetric lines, and were therefore clearly carried out in passing by an individual who was not concerned with the lines' metre. In the examples in *Jehan Paulus*, by contrast, all seven lines fit the alexandrine metre, such that the intervention of a more committed *remanieur* would have been required to ensure that the modified forms still retained the desired metre.

The evidence presented in this chapter has therefore reinforced Munk Olsen's warning that attribution cannot be made on this feature alone, and it has also suggested that the non-use of *mon* etc. before feminine vowel-initial nouns in the *J-corpus*, *Guillaume*, and *Robert* is not especially surprising, given their temporal and codicological context. Nevertheless, it has also strengthened the utility of this feature as a means of negative attribution: the metrically consistent use of *mon* etc. before feminine vowel-initial nouns in *Jehan Paulus* and *Richard* does appear to be enough to exclude them from the core of Jehan's *œuvre*, at least in their current form. The undeniable proximity of these texts to the *J-corpus*, based on the other features that we have examined so far in this study, must instead be taken as evidence that these are examples of earlier texts by Jehan which were significantly, albeit not entirely, reworked by later poets.

## Building the Case for the Shared-Author Hypothesis

The linguistic evidence examined in this chapter allows us to nuance the shared-author hypothesis by distinguishing between those texts that merit direct attribution to Jehan, and those that are later reworkings of earlier poems by the same author. The consistency in the use of elided *m'*, *t'*, and *s'* before feminine vowel-initial nouns in the *J-corpus* offers an additional unifying characteristic to the corpus, while the strict adherence of *Guillaume* and *Robert* to this practice makes them clear candidates for direct inclusion in Jehan's *œuvre*. *Jehan Paulus* and *Richard*, meanwhile, display a fluctuating usage which mixes elided forms with the analogical use of *mon*, *ton*, and *son* which later becomes the norm. Because these analogical forms do not produce hypermetric lines, they point towards the intervention of one or more later *remanieurs* who updated what they saw as archaic usage in an earlier poem by Jehan. This observation is crucial to the shared-author hypothesis, in that it identifies *Jehan Paulus* and *Richard* as echoes of texts originally composed by Jehan which were then significantly, though not entirely, reworked by a later poet with different linguistic preferences.

## Methodological Contribution

In this chapter, I have shown how digital methods can offer a rigorous, falsifiable base from which to critically interrogate the assertions of traditional philology. Through the examination of reference corpora which are contemporary with our target corpus, I have demonstrated that features previously considered unique to a specific author may in fact be consistent with broader temporal and dialectal norms, and that awareness of such norms can prevent the misattribution of authorship based solely on the identification of widespread linguistic features. Moreover, by examining the metrical consistency of apparent modifications, I have demonstrated the value of such an approach in distinguishing between a text that has been reworked by a *remanieur*, and one which may simply have been modified by a copyist with a preference for a given linguistic feature.

## VERSIFICATION

FROM THE EARLIEST attribution of texts to Jehan, the primary motivating factor for doing so appears to have been a sense amongst scholars that the texts could be connected “wegen der ihnen eignen Einfalt des Vortrags und des unvollkommenen Sprach- und Reimvermögens” (Gröber 1888–1902, 2: 929).<sup>1</sup> Scholars have taken issue with Jehan’s apparent “médiocrité technique” when it comes to versification (Munk Olsen 1978, cxxxvi),<sup>2</sup> using the apparent ubiquity of such failings across the *J-corpus* as grounds to suggest that the texts share a single author.

---

1 “because of the inherent simplicity of their delivery and their imperfect command of language and rhyme.”

2 “technical mediocrity.”

---

**ABSTRACT** This chapter examines Jehan’s versificatory practices in the context of the traditional scholarly view of his apparently mediocre poetic capabilities. Rather than engaging in unhelpful debates around value, I instead take the key feature which has previously been presented to critique Jehan’s work, namely the frequency of so-called “easy” rhymes, alongside other features such as the rates of use of masculine and feminine rhymes. By comparing the *J-corpus* to the generically comparable *A-* and *R-corpora*, I show that although the *J-corpus* is internally variable for both features, there is a coherence to this variability which differentiates it from the reference corpora. The most significant methodological contributions of this chapter are the introduction of the concept of a *rhyme coefficient*, a measure of how likely a word is to appear in rhyming position compared to elsewhere in the line, and our analysis of admixtures, that is pairs of words which frequently appear together in rhyming position. These techniques allow us to extract the clearest evidence so far of an identifiable authorial fingerprint, revealing that there were specific stock words which Jehan particularly favoured at the end of a line, especially when compared to the generically and formally comparable *A-corpus*. The examination of admixtures, meanwhile, particularly reinforces the shared-author hypothesis for *Guillaume*, *Robert*, *Richard*, and *Jehan Paulus*, the four additional texts which we have associated with Jehan’s *œuvre*.

---

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 overuse 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. 205–10; *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 admit, 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."<sup>3</sup> Within the *A-corpus*, for instance, we may note that stanzas of more than four lines are found in *Confesseur* (five-line stanza vv. 358–62), *Vision* (five-line stanza vv. 566–70), *Thibaut* (six-line stanza vv. 1053–58), and *Vieillards* (five-line stanza vv. 24–28). None of these texts have so far stood out as potential candidates for Jehan's *œuvre*, revealing the difficulty of building a case for shared authorship on this feature. Ultimately if, across similar surviving texts from the period, "la plupart des spécimens [...] ne se distinguent guère par une technique souveraine" (Munk Olsen 1963b, 157),<sup>4</sup> 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.

Indeed, 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 is never consistent in quality, as they hone and develop their craft across their career, or fall from earlier creative heights. The avenues pursued by previous scholarship regarding Jehan's versification therefore offer little evidence 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 the existing claims concerning Jehan's versificatory tendencies, as well as pursue new avenues. Combined, these can tell us much more than previous scholarship has been able to determine about the extent to which the versificatory practices of the *J-corpus* unite it as a coherent group, distinct from texts found elsewhere.

<sup>3</sup> "[a]ll these irregularities are found, often to an even greater degree, in other poems of the period, written in the same metrical form."

<sup>4</sup> "most specimens [...] are hardly distinguished by their masterful technique."

## 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.<sup>5</sup> 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 poems, which he claims is "la norme" (the norm) for texts of this period. As we shall see, the frequency with which Jehan uses feminine rhymes is not identical to that observed in texts of a comparable type, but it is within the expected range for the fourteenth century, given Lote's (1949–1996, 2:111–24) observation that the rate of use of feminine rhymes increases at a roughly even rate from near to zero in the earliest surviving poems in French, to near-parity with masculine rhymes by the end of the fifteenth century.

However, this rate of use of feminine rhymes is not without variability across the *J-corpus*, as shown in Figure 5. 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 recognizing 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%).

One may choose to hypothesize that some explanation for the increased use of feminine rhymes in these texts may result from the fact that, except for *Pain*, all of these texts displaying high rates of feminine rhymes centre

<sup>5</sup> 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 represented through regular expressions, namely that at least two lines of a stanza must satisfy `es?|[a-zéäëïöüÿ]{2,}ent$`, and no line in the stanza can satisfy `ment$|ant$|^[a-z]i?ent$|é$`. 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. For more details on the use of regular expressions, see p. 19.

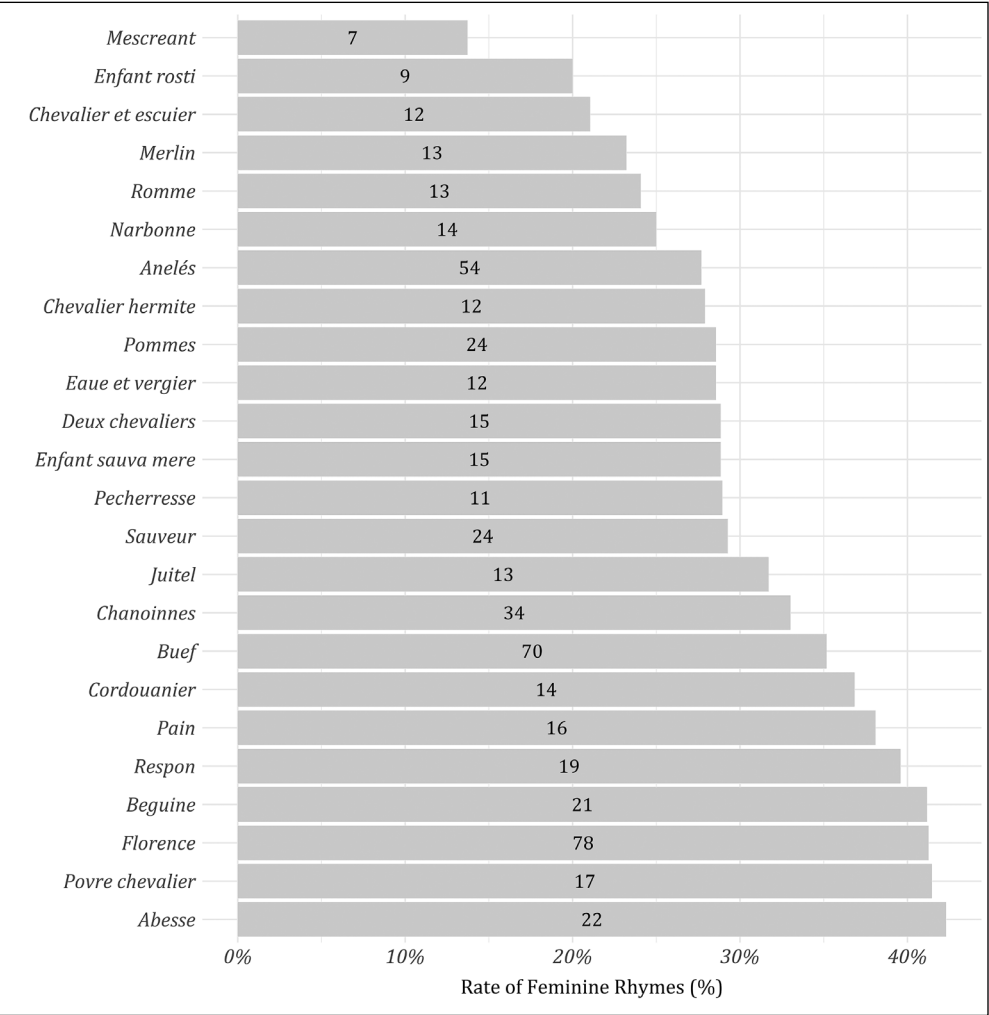


Figure 5. Rates of use (x-axis) and raw numbers (figures) of stanzas containing feminine rhymes for each text of the *J-corpus*.

around the experience of women,<sup>6</sup> 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. 114–15). 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

**6** In *Povre chevalier*, it is the actions of the knight’s wife and of Mary which form the crux of the story.

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.

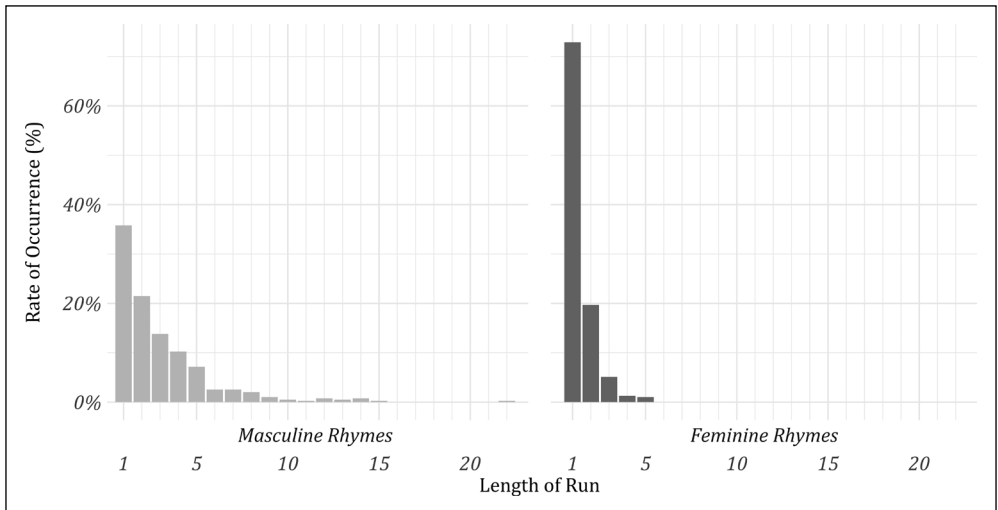


Figure 6. Rates of occurrence for runs of stanzas containing masculine and feminine rhymes across the *J-corpus*.

Munk Olsen (1978, cxxxix) argues that within Jehan's texts, monorhyme stanzas containing masculine rhymes appear alongside those containing feminine rhymes "sans principe fixe."<sup>7</sup> Overall, 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 6, we note that the appearance of stanzas

<sup>7</sup> "without any fixed principle."

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%). That is, stanzas containing feminine rhymes are about twice as likely to appear on their own, flanked by two stanzas of masculine rhymes, than the inverse. This is perhaps at least somewhat due to 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 two stanzas to be the most frequent type for masculine rhymes, with masculine and feminine rhymes alternating on a two-to-one basis.

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 five (twice in *Florence*, once in *Buef*, and once in *Abesse*), while for masculine rhymes it is twenty-two (*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 (five), we find that only 770 stanzas of a total of 1,172 fit these criteria, putting the stanzas with masculine rhymes at a little under 50% greater than the number of stanzas containing feminine rhymes (539). For the most part, then, the 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 *A-corpus* of narrative alexandrine monorhyme poems which have hitherto not been attributed to Jehan, 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.

From Figure 7, we observe that the median rate of use of feminine rhymes in the *J-corpus* (28.9%) is relatively comparable to that of the *A-corpus* (31.8%), suggesting that there is an expected rate for alexandrine monorhyme quatrains from this period. The range, however, is vastly greater. The *A-corpus* contains one lower outlier (*Troyes*, 0%) and two upper outliers

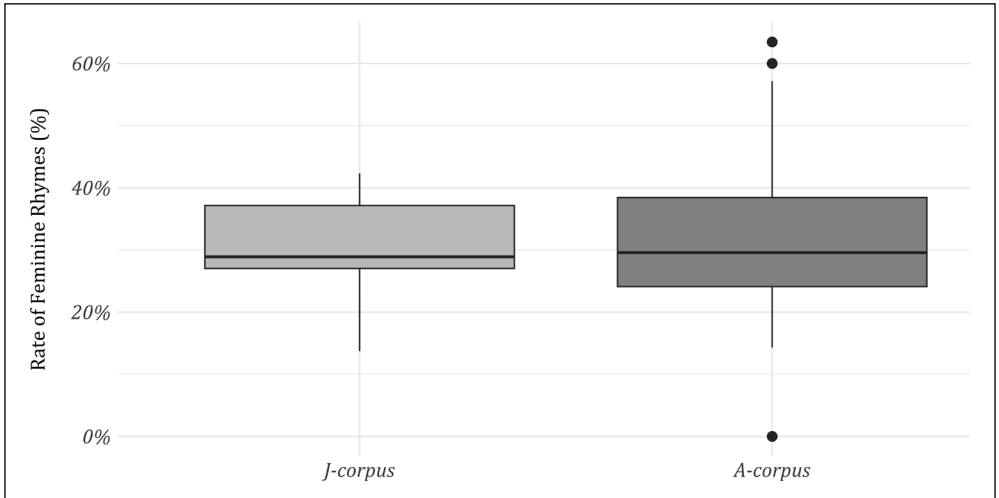


Figure 7. Box and whisker plots for the rate of use of feminine rhymes in the *J-* and *A-corpora*.

(*Dieu d'Amours*, 60%; *Quatre sereurs*, 63.5%), and a further four texts display rates greater than the highest rate of the *J-corpus* (*Abesse*, 42.3%), namely *Confesseur* (44.1%), *Barbe* (44.6%), *Thomas* (47.4%), and *Yves* (57.1%). As a result, while it is possible to dismiss *Yves* and *Dieu d'Amours* on account of their brevity, at just seven and five stanzas respectively, the remaining evidence provides coherence to the *J-corpus* by suggesting that Jehan's work is much less variable than other comparable texts.

The rate of use of feminine rhymes, then, appears to be broadly comparable across texts of this kind, but it is at least in part also a feature of authorial preference, since variability is possible. It is therefore all the more notable, looking at the individual rates of the *A-corpus* in Figure 8, that *Guillaume* and *Robert*, the two texts which we have thus far hypothesized to be the work of Jehan, as well as *Jehan Paulus* and *Richard*, the apparent echoes of earlier works by the same author, 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%, 28.7%, 29.0%, and 29.1% respectively. This overlap heightens the association between these texts, and reinforces the shared-author hypothesis for these texts alongside the *J-corpus*. Of course, the presence of similar rates for *Basile*, *Alexis*, *Grégoire*, *Sebastien*, *Vieillards*, and *Dieudonné* means that this evidence on its own is not enough to confirm attribution, and but it does still contribute to the more holistic evidence which we are building over the course of this study.

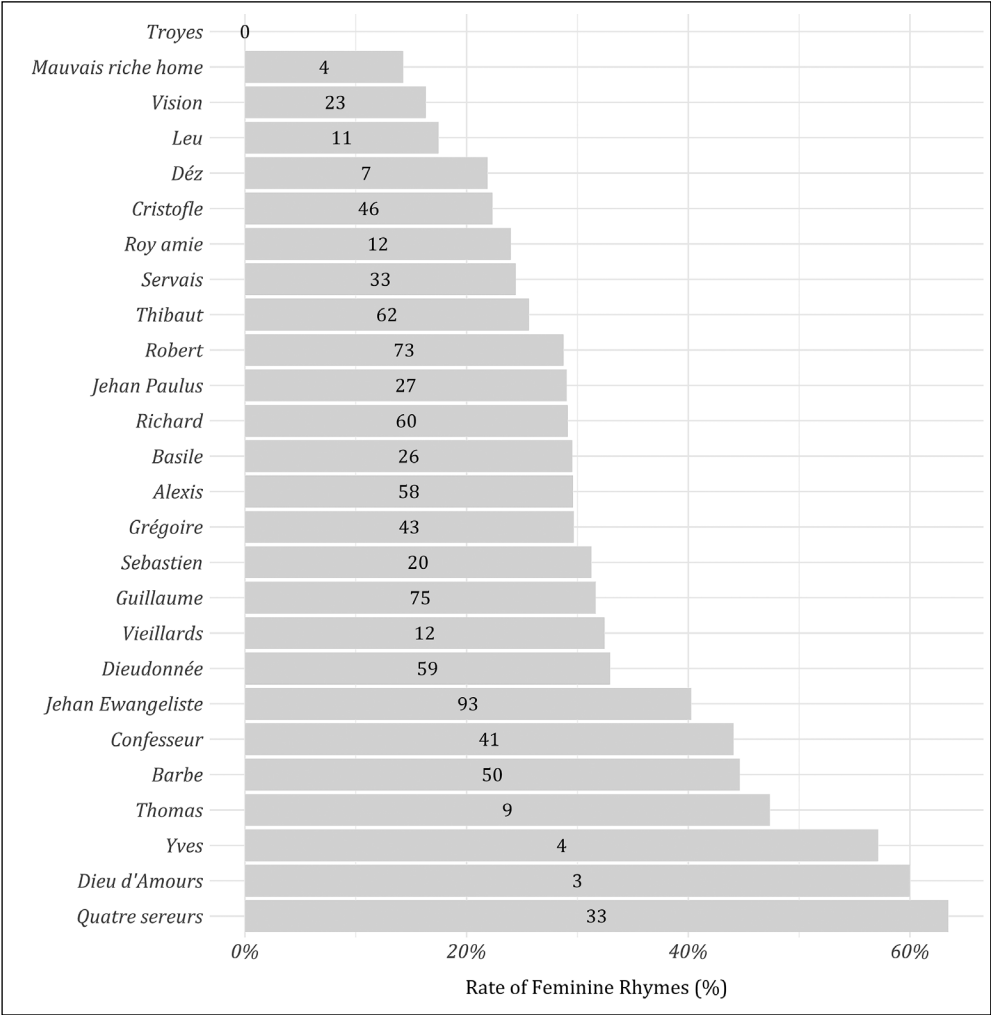


Figure 8. Rates of use (x-axis) and raw numbers (figures) of stanzas containing feminine rhymes for each text of the *A-corpus*.

Returning also to the “runs” of stanzas containing masculine versus feminine rhymes, as represented in Figure 9, we observe a similar pattern in the *A-corpus* to that seen for the *J-corpus*, namely that runs of masculine rhymes are often much longer than runs of feminine rhymes. The longest run of feminine rhymes in the corpus is seven (one each in *Barbe* and *Jehan Ewangeliste*), while the longest run of masculine rhymes is twenty-six (in *Leu*), with runs of twenty-two and twenty-four stanzas also found in *Alexis* and *Vision* respectively. *Guillaume*, *Robert*, *Richard*, and *Jehan Paulus*, meanwhile, align with the broader trends of the *J-corpus*, with one ten-stanza run

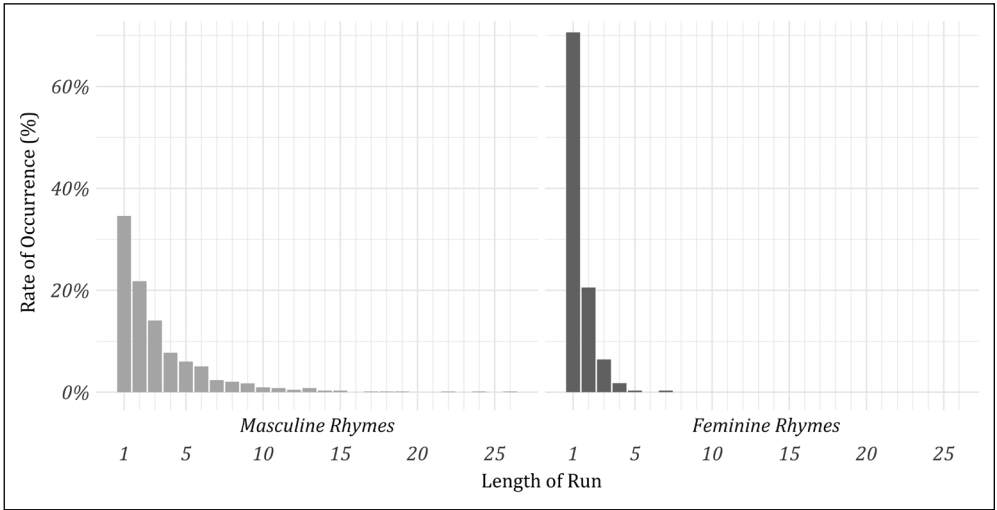


Figure 9. Rates of occurrence for runs of stanzas containing masculine and feminine rhymes in the *A-corpus*.

of masculine rhymes appearing in *Guillaume*, two twelve-stanza runs in *Robert*, one nineteen-stanza run in *Richard*, and one ten-line run in *Jehan Paulus*.

As with the *J-corpus*, by excluding those runs of stanzas containing masculine rhymes which exceed a length of five, the number of masculine stanzas in *Guillaume*, *Robert*, *Richard*, and *Jehan Paulus* drops from 555 to 347, and is approximately 50% greater than the number of feminine stanzas (235), therefore suggesting that for these texts, 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.

To conclude our section on feminine rhymes, we find that the usage observed for the *J-corpus* is arguably close to typical, as previous scholarship has suggested, but still displays some variability within these rates. Some level of variation in the use of feminine rhymes appears to be possible both within the *œuvre* of one single author and between the work of different authors, and while we are occasionally able to explain this variation by considering a text's content, it is frequently impossible to do so. Evidence of far more variability within other narrative texts written in monorhyme alexandrine quatrains, however, supports the collective attribution of the *J-corpus* to a single author, as well as reinforcing the shared-author hypothesis for *Guillaume*, *Robert*, *Richard*, and *Jehan Paulus*, and suggesting that the *remanieur(s)* of the latter two texts did not make especially significant changes to Jehan's original versification.

## 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), “easy” rhymes, 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).<sup>8</sup> Morawski (1939, 334) defines this class of “easy” rhymes 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 it is 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 solely to reach a clearer understanding of the idiosyncratic uses of such rhymes that, so it has been claimed, serve to unite the corpus, and in so doing to further interrogate the hypothesis that the texts share an author.

Our focus here will be on two specific types of “easy” rhyme which are claimed 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 valetton, si l’amerent forment.  
Li preudon n’en ot plus en trestot son vivant.

(*Romme*, vv. 13–16)

---

<sup>8</sup> “distinguish the good poet from the hopelessly bad versifier.”

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.<sup>9</sup> 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.<sup>10</sup>

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. 769–72)

Through examination of the uses of "easy" rhymes in Jehan's corpus of texts, as represented in Figure 10, 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 11.3% for the 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 ( $-0.37247$ ). A Pearson's product-moment correlation test suggests that the correlation is not quite

**9** 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 each stanza must match to be counted is therefore `an[tsz]$|ment$`.

**10** 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.

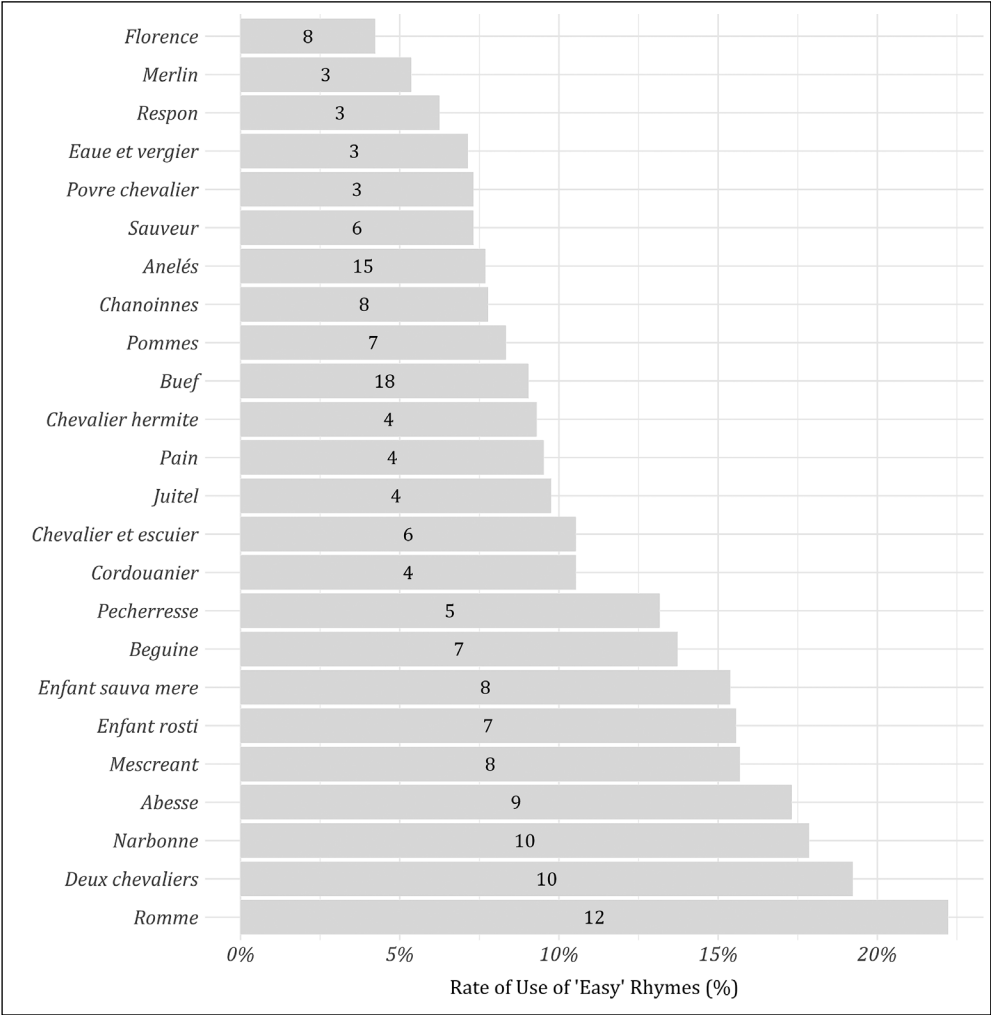


Figure 10. Rates of use (x-axis) and raw numbers (figures) of stanzas containing “easy” rhymes in the *J-corpus*.

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.<sup>11</sup>

11 No such significant correlation appears between the use of feminine rhymes and text length.

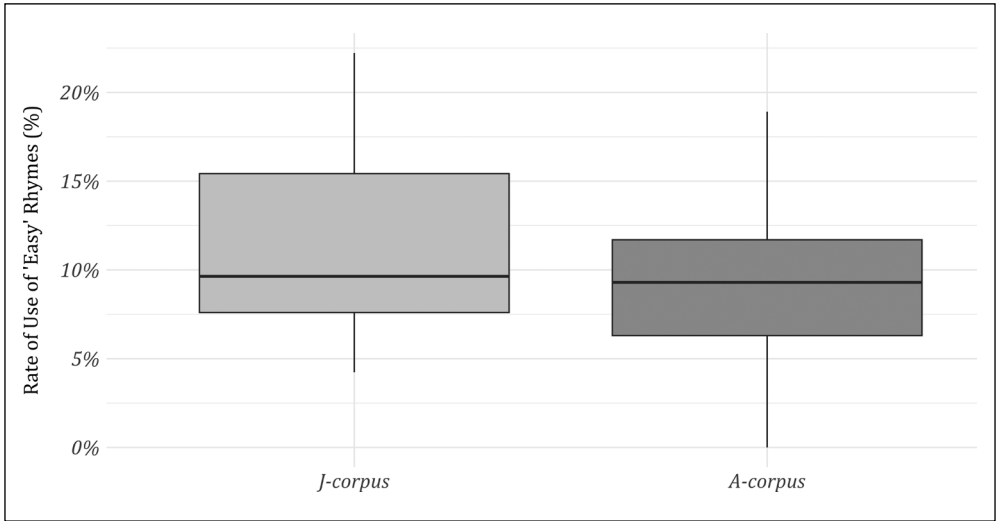


Figure 11. Box and whisker plots for the rate of use of “easy” rhymes in the *J-* and *A-corpus*.

When we compare the rates of use of “easy” rhymes in the *J-corpus* to the *A-corpus* containing narrative poems in alexandrine monorhyme quatrains which are not usually attributed to Jehan, as represented in Figure 11, we observe that the median rate of the *A-corpus* (9.30%) is very comparable to that of the *J-corpus* (9.64%), but also that the range of the data is lower than that of the *J-corpus*, while the interquartile range is wider (*J*: 7.8%; *A*: 5.4%). All of this serves to reinforce the suggestion that in many cases, if not on average, Jehan’s texts display a greater rate of use of “easy” rhymes than other, comparable poems, and across a broader spread.

Looking at the individual rates for each of the texts in Figure 12, we again observe similarities with Figure 10, and note that *Guillaume*, *Robert*, *Richard*, and *Jehan Paulus*, the four texts to which we have extended the shared-author hypothesis, all fit comfortably within the range of the *J-corpus*. Beyond this, there are also differences between the corpora in that, even excluding the fragmentary *Dieu d’Amours*, there are three texts in the *A-corpus* which display lower rates of “easy” rhymes than the lowest rate in the *J-corpus* (namely *Florence*, at 4.2%), and only two which exceed a rate of 15% (*Troyes* and *Vieillards*), compared to seven in the comparably-sized *J-corpus*. In the *A-corpus*, there is also no correlation between stanza count and rate of use of “easy” rhymes which even approaches statistical significance ( $p = 0.8718$ ), such that there are clearly some differences between the preferences for “easy” rhymes expressed in the *J-corpus* and those of the *A-corpus*.

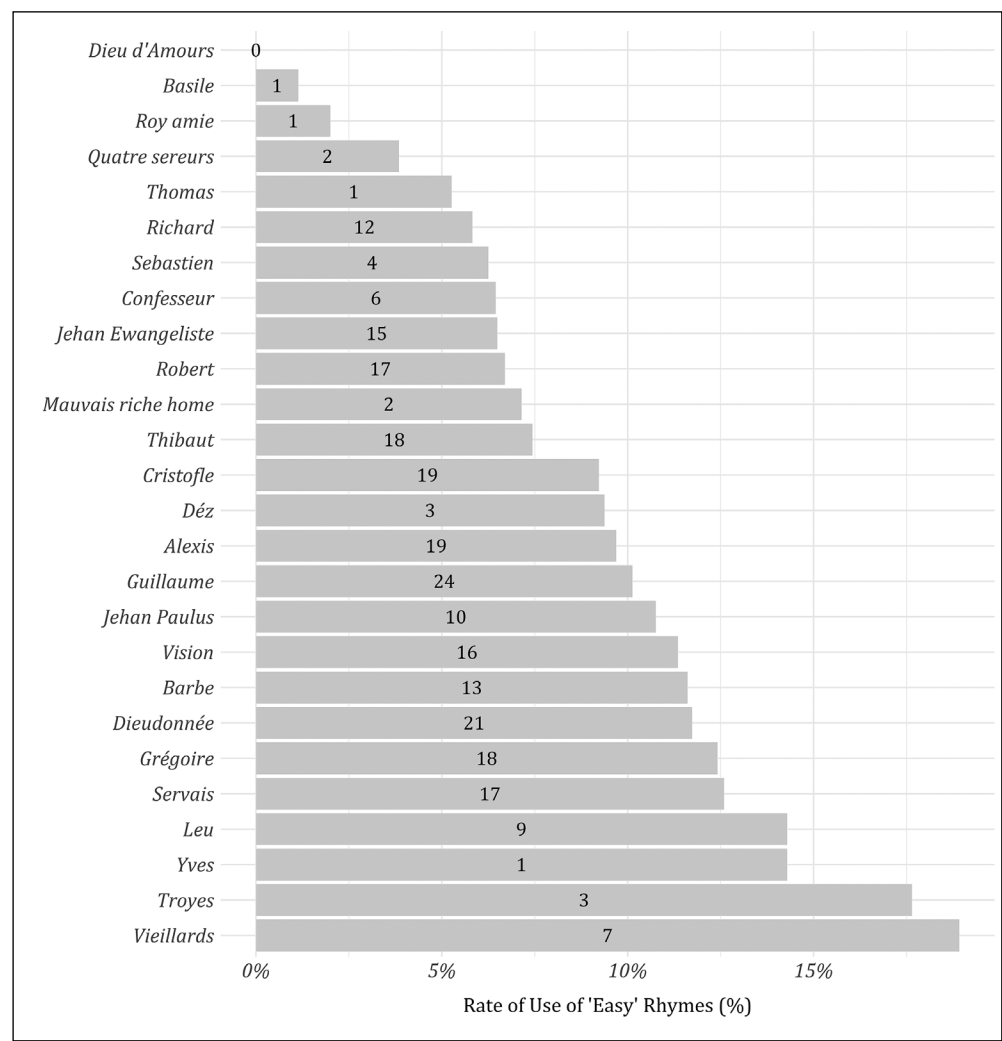


Figure 12. Rates of use (x-axis) and raw numbers (figures) of stanzas containing “easy” rhymes in the *A-corpus*.

Overall, then, we are able identify certain features of Jehan’s versificatory preferences through examination of “easy” rhymes, in particular his willingness to allow higher rates in some of his texts, especially the shorter ones. However, an additional statistical test here reveals that it is not straightforwardly true that Jehan simply makes greater use of “easy” rhymes than those writing comparable works, as the existing scholarship claims. A Welch Two Sample t-test here produces a result ( $p = 0.09249$ ) that is not quite sufficient to suggest that there is a statistically significant difference

between the two corpora for this feature, and this remains true even if we exclude *Guillaume*, *Robert*, *Richard*, and *Jehan Paulus* ( $p = 0.1369$ ) from the *A-corpus*, or include them alongside the *J-corpus* ( $p = 0.2032$ ). That is, while such a suggestion appears plausible, the two ranges are sufficiently similar that we are unable to conclude with certainty that their difference results from anything other than random variation, and thus attributed authorship alone cannot account for variation in this feature.

Following these observations in the *A-corpus*, and as a final point of analysis in this section, it is worthwhile comparing the *J-* and *A-corpora* to the *R-corpus*, containing the complete surviving poetic works of Rutebeuf. In doing so, we can determine whether the rate of use of “easy” rhymes is an apparently standard feature of narrative poems written in alexandrine monorhyme quatrains, when compared to Rutebeuf’s *œuvre* which is more diverse in form.<sup>12</sup>

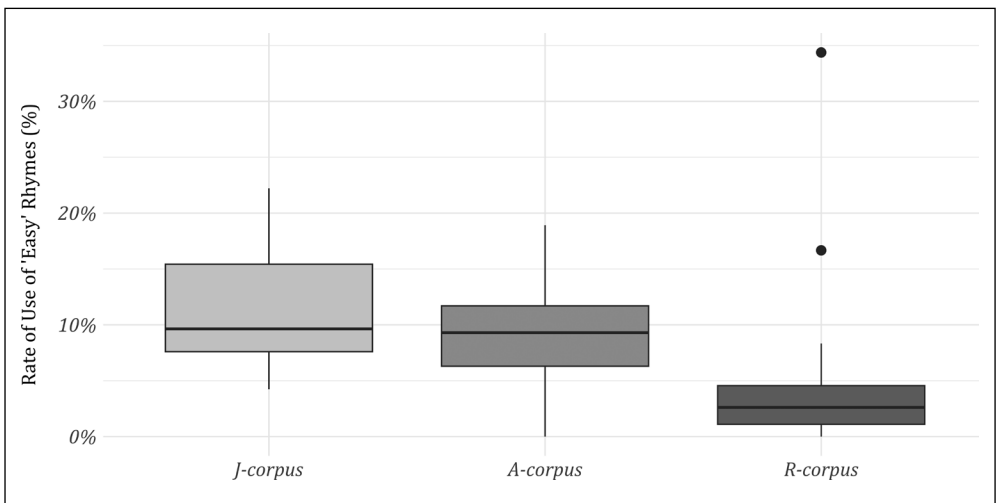


Figure 13. Box and whisker plots for the rate of use of “easy” rhymes in the *J-*, *A-*, and *R-corpora*.

The data for this comparison are presented in Figure 13. In this case, for the *R-corpus*, we note that the median use of “easy” rhymes is much lower than that observed for the *J-* and *A-corpora*, with thirteen texts of the fifty-six in the *R-corpus* not containing even a single instance of an “easy”

<sup>12</sup> Given this formal variety, when calculating the rate of “easy” rhymes for the *R-corpus*, it is necessary to consider the texts on a line-by-line basis, since we are often not able to divide the text into monorhyme stanzas.

rhyme. Notably, this includes two of Rutebeuf's poems written in alexandrine monorhyme quatrains, namely *Li diz des Cordeliers* and *Des Jacobins* (neither of which are strictly narrative in their content). 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 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 eight-line stanzas. The relatively high rate in *De Bricheuer* (16.7%), meanwhile, appears more straightforwardly to be a function of the brevity of this poem, with "easy" rhymes appearing in just four of the poem's twenty-four lines, all rhyming with each other as one of the two rhymes in an eight-line alternating stanza.

Overall, the *R-corpus* contains significantly lower rates than the *J-corpus* ( $p = 0.00000$ ) and the *A-corpus* ( $p = 0.00004$ ). This observation allows us to reframe Munk Olsen's claim that, for poems written in alexandrine monorhyme quatrains, "la plupart des spécimens [...] ne se distinguent guère par une technique souveraine" (1963b, 157).<sup>13</sup> Setting aside his value judgement, it does nevertheless appear to be true that Jehan was in good company alongside the other authors of narrative poems written in alexandrine monorhyme quatrains, in that he displays a notable preference for rhymes ending in *-ment* and *-ant*. What is equally clear, however, is that this feature alone, while allowing us to identify outliers like *Basile* which are unlikely to be the work of Jehan, cannot enable us to claim that Jehan's use of easy rhymes somehow marks him out as a "versificateur irrémédiablement mauvais" (Munk Olsen 1963b, 158)<sup>14</sup> relative to his peers.

## Word Frequency in Rhyming Position

Alongside his claims about the quality of Jehan's poetry, Munk Olsen (1963b, 160) suggests a further element of versification which may prove effective in considering the shared-author hypothesis for Jehan's poems. It may be fruitful, Munk Olsen argues, to consider *associations des rimes* or *rhyme associations*, 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

<sup>13</sup> "most specimens [...] are hardly distinguished by their masterful technique."

<sup>14</sup> "hopelessly bad versifier."

Olsen's method appears promising, since it allows him to compare his corpus of Jehan's texts to *Alexis* (included in our *A-corpus*) through analysis of the rhymes *-on(s)* and *-is*.<sup>15</sup> 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.<sup>16</sup> We then divide the former rate by the latter to determine what I term the *rhyme coefficient*, that is, how much more or less likely a given word is to appear in rhyming position than elsewhere in the line. This method allows us to make a critical distinction between those words which appear frequently due to the simple fact that they are often used across the corpus, because of authorial preference or narrative necessity, and those words which could be described as "stock" rhyming words, used to complete a rhyme, but rare elsewhere.

In this section of our analysis, we do not focus on the word forms that appear at the end of lines, rather their associated lemmata, that is, the arbitrarily determined "dictionary" form of a lexeme, as given 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 ten 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 single authorship is more assured. Across the fifty-six texts of the *R-corpus*, we observe some forty-nine lemmata which match our selection

---

**15** The corpus Munk Olsen uses for this analysis is shorter than the 1978 corpus on which he eventually settles.

**16** 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*), while the non-final rate is calculated using the total number of non-final tokens in the corpus (52,636).

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, *longument* (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 fifty 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 our approach has significant validity as a means of identifying the individual versificatory “fingerprint” of a given author.

Applying this analysis across the *J-corpus*, we find several trends of note. Table 32 shows the top twenty lemmata across the corpus, when ordered by descending rhyme coefficient. First, 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 (their rhyme coefficient is therefore infinite). We also note that many lemmata with a high rhyme coefficient, including *certainement*, *vraiment*, *commandement*, and *sauvement*, are examples of the so-called “easy” rhymes discussed above.<sup>17</sup>

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.<sup>18</sup> 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

---

**17** Across all lemmata which appear on more than ten occasions, we find a further twenty 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), *longument* (6.9), *couvent* (6.4), *erramment* (5.8), *avant* (4.2), *brïement* (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 fifteen of the forty-nine forms with a rhyme coefficient above 1 counting as “easy” rhymes, we can further resist the claims of previous scholarship that an abundance of such forms has anything to do with poor versificatory practice.

**18** 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 focused on the forms with the highest rhyme coefficients.

Table 32. Counts and rates of lemmata which appear on ten or more occasions in the *J-corpus*, limited to the top twenty results.

| Lemma               | Line-final count | Line-final rate | Non-final count | Non-final rate | Rhyme coefficient |
|---------------------|------------------|-----------------|-----------------|----------------|-------------------|
| <i>certainement</i> | 13               | 0.19%           | 0               | 0.00%          | —                 |
| <i>envie</i>        | 11               | 0.16%           | 0               | 0.00%          | —                 |
| <i>demeuree</i>     | 10               | 0.15%           | 0               | 0.00%          | —                 |
| <i>contree</i>      | 12               | 0.18%           | 1               | 0.00%          | 92.5              |
| <i>plevir</i>       | 12               | 0.18%           | 1               | 0.00%          | 92.5              |
| <i>vraiment</i>     | 12               | 0.18%           | 1               | 0.00%          | 92.5              |
| <i>arreter</i>      | 11               | 0.16%           | 1               | 0.00%          | 84.8              |
| <i>partie</i>       | 21               | 0.31%           | 2               | 0.00%          | 80.9              |
| <i>commandement</i> | 10               | 0.15%           | 1               | 0.00%          | 77.1              |
| <i>appareiller</i>  | 9                | 0.13%           | 1               | 0.00%          | 69.4              |
| <i>usage</i>        | 9                | 0.13%           | 1               | 0.00%          | 69.4              |
| <i>tenter</i>       | 13               | 0.19%           | 2               | 0.00%          | 50.1              |
| <i>pensee</i>       | 19               | 0.28%           | 3               | 0.01%          | 48.8              |
| <i>endurer</i>      | 11               | 0.16%           | 2               | 0.00%          | 42.4              |
| <i>chiere</i>       | 10               | 0.15%           | 2               | 0.00%          | 38.5              |
| <i>corage</i>       | 20               | 0.29%           | 4               | 0.01%          | 38.5              |
| <i>esbahir</i>      | 9                | 0.13%           | 2               | 0.00%          | 34.7              |
| <i>debonnaire</i>   | 8                | 0.12%           | 2               | 0.00%          | 30.8              |
| <i>matiere</i>      | 8                | 0.12%           | 2               | 0.00%          | 30.8              |
| <i>sauvement</i>    | 8                | 0.12%           | 2               | 0.00%          | 30.8              |

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.

The lack of much significant overlap between the forms with high coefficients in 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 *A-corpus*, it is possible to identify trends which assist us in attributing the authorship of the *J-corpus* to a single author. Table 33 displays the top twenty lemmata for the *A-corpus*, when ordered by rhyme coefficient, allowing us to make direct comparisons to the texts of the *J-corpus*.

Table 33. Counts and rates of lemmata which appear on ten or more occasions in the *A-corpus*, limited to the top twenty results.

| Lemma                 | Line-final count | Line-final rate | Non-final count | Non-final rate | Rhyme coefficient |
|-----------------------|------------------|-----------------|-----------------|----------------|-------------------|
| <i>agreer</i>         | 13               | 0.11%           | 0               | 0.00%          | —                 |
| <i>embraser</i>       | 12               | 0.10%           | 0               | 0.00%          | —                 |
| <i>demeuree</i>       | 11               | 0.09%           | 0               | 0.00%          | —                 |
| <i>droiturier</i>     | 10               | 0.09%           | 0               | 0.00%          | —                 |
| <i>passion</i>        | 11               | 0.09%           | 1               | 0.00%          | 84.9              |
| <i>debonnairement</i> | 10               | 0.09%           | 1               | 0.00%          | 77.2              |
| <i>lumiere</i>        | 10               | 0.09%           | 1               | 0.00%          | 77.2              |
| <i>renommee</i>       | 10               | 0.09%           | 1               | 0.00%          | 77.2              |
| <i>targer</i>         | 10               | 0.09%           | 1               | 0.00%          | 77.2              |
| <i>oublier</i>        | 9                | 0.08%           | 1               | 0.00%          | 69.5              |
| <i>honorer</i>        | 17               | 0.14%           | 2               | 0.00%          | 65.6              |
| <i>contraire</i>      | 10               | 0.09%           | 2               | 0.00%          | 38.6              |
| <i>droiture</i>       | 10               | 0.09%           | 2               | 0.00%          | 38.6              |
| <i>vis</i>            | 10               | 0.09%           | 2               | 0.00%          | 38.6              |
| <i>folie</i>          | 19               | 0.16%           | 4               | 0.00%          | 36.7              |
| <i>retraire</i>       | 19               | 0.16%           | 4               | 0.00%          | 36.7              |
| <i>saisir</i>         | 14               | 0.12%           | 3               | 0.00%          | 36.0              |
| <i>prouver</i>        | 9                | 0.08%           | 2               | 0.00%          | 34.7              |
| <i>remüer</i>         | 9                | 0.08%           | 2               | 0.00%          | 34.7              |
| <i>courage</i>        | 16               | 0.14%           | 4               | 0.00%          | 30.9              |

The first main observation of note is that only *demeuree* appears with an infinite rhyme coefficient across both the *J-* and *A-corpora*, suggesting that this lemma is particularly likely to appear in rhyming position across the work of multiple authors. The lack of overlap on other words, meanwhile, suggests that authorial preference may be more at play in these cases. Broader comparison reinforces this suggestion further, since several lemmata have much higher rhyme coefficients in the *J-corpus* than in the *A-corpus*, for example *commandement* (*J-corpus*: 77.1; *A-corpus*: 7.7), *endurer* (42.4 to 4.9), *con-tree* (92.5 to 11.0), and *tenter* (50.1 to 7.7). Looking in the other direction is similarly fruitful, with *mari* (1.9 to 11.6), *aimer* (1.9 to 9.9), *defendre* (2.8 to 14.2), and *tendre* (5.1 to 25.7) all appearing with markedly greater rhyme coefficients in the *A-corpus* than in those texts previously attributed to Jehan.

It is therefore clear that, for certain words, the rhyme coefficient differs a great deal between the *J-* and *A-corpora*, suggesting a role played by authorial preferences and therefore supporting the shared-author hypothesis for the *J-corpus*.<sup>19</sup> This does not appear to be the case for every word that appears at the rhyme, however, 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 *A-corpus*. Examining those lemmata within the *J-corpus* which appear at the rhyme in at least one instance and at least ten times anywhere in the line across each of the *J-* and *A-corpora*, we find a statistically significant, weakly positive correlation (0.42489,  $p = 0$ ) between the rhyme coefficients for individual lemmata. We can say with some certainty, then, that apart from 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 *A-corpus*.

Producing a linear regression model for these data, however, reveals that the link between these two corpora is not straightforwardly 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 = 3.0553, gradient = 0.2910) 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 32.16 in the *A-corpus*. While we can therefore expect certain lemmata to have a greater rhyme coefficient across both the *J-* and *A-corpora*, Jehan reserves certain words for use solely, or primarily, at the ends of lines to a much greater extent than is observed in the more heterogeneous *A-corpus* where multiple competing preferences are in play from the corpus' various authors. This observation thus reinforces the suggestion that the *J-corpus* is a single-author corpus, whereas the *A-corpus* is not.

Figure 14 maps rhyme coefficients in both corpora as a scatter plot, along with the calculated regression line (dashed line), and a line indicating

---

**19** We note, of course, at this stage, that the *A-corpus* contains *Guillaume, Robert, Richard*, and *Jehan Paulus*, to which we have extended the shared-author hypothesis. The methods discussed in this section lend themselves best to examination of large-scale corpora, rather than individual texts, since the latter are likely to be influenced in their word choices by their content, and so it has therefore not been possible to apply this analysis to this smaller sub-group of texts. The analysis here therefore concerns the extent to which the *J-corpus* displays a coherence which the *A-corpus* does not, regardless of the identity of the authors of the *A-corpus* texts, and as such it is not problematic that the *A-corpus* contains texts which appear to also be the work of Jehan.

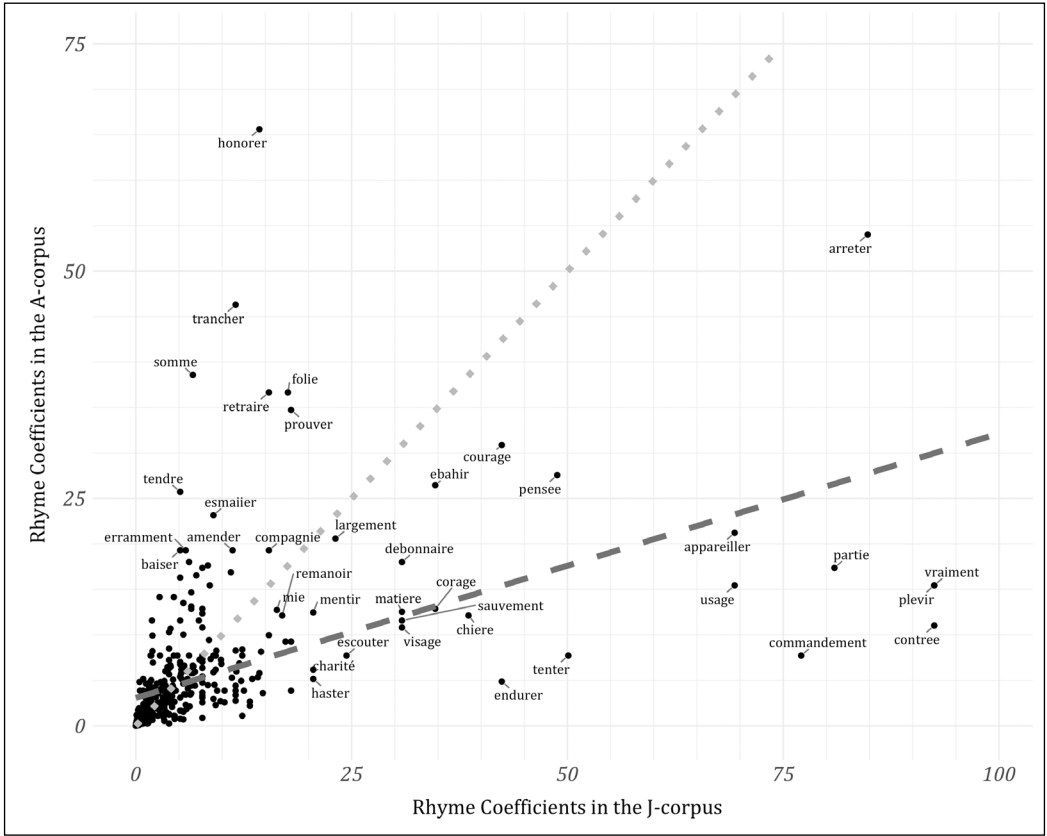


Figure 14. Scatter plot for the rhyme coefficients of lemmata found in rhyming position in the *J*- and *A*-corpora.

a putative one-to-one correlation (dotted line).<sup>20</sup> 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 *A*-corpora, including *largement* and *ebahir*. 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 broader contemporary expectations of style for poems written in alexandrine monorhyme quatrains. Alongside this, we also get a clearer picture of the outlier status of lemmata such as

**20** Several 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.

*endurer*, *tenter*, *commandement*, and *contree*, which have much higher rhyme coefficients in the *J-corpus* than in the *A-corpus*, and which Jehan therefore uses more frequently at the rhyme than comparable poets, as well as *tendre*, *somme*, *trancher*, and *honorer*, for which the opposite is true. Lemmata such as *matiere*, *corage*, and *appareiller*, meanwhile, stick closely to the calculated regression line, indicating that while they are common at the rhyme in both the *J-* and *A-corpora*, they are still much more likely to appear in rhyming position in Jehan's poems.

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, authorial identity is a significant factor in determining rhyme coefficients. The examination of this feature has therefore added to the overall cohesiveness of the *J-corpus*, strengthening the shared-author hypothesis.

## 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 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, 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 at the rhyme on five 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.

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 five times alongside a context lemma and twice elsewhere has a ratio of 2.5:1 (given simply as 2.5). A lemma is said to exist in a prominent admixture with another lemma if this ratio is greater than 1, meaning that the lemma appears alongside its associated context lemma more frequently than it does elsewhere.

There are thirty-four pairs of lemmata which satisfy these criteria within the *J-corpus*, suggesting that they are pairings which the Jehan used together on a particularly frequent basis. The lemma *faire* has four associated lemmata which appear in prominent admixture with it (7 with *debonnaire*,

7 with *retraire*, 5 with *contraire*, and 5 with *exemplaire*),<sup>21</sup> 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 form at the rhyme is the infinitive for the first form, or the third person singular in the present tense or the past participle for the last three forms.

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 can also be said 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-retraire* (both 1.67), *chemin-fin* (2 and 1.33), *manière-chiere* (2 and 1.5), and *mere-pere* (1.75 and 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.

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* is also a context lemma for *clair* (5) and takes *amer* (1.75) as a context lemma. Of the remaining examples, we find four lemmata with *vie* as a context lemma (*compagnie*, 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* (*compagnie*, 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.

To identify how unusual, or otherwise, these admixtures are, we must apply the same approach to the *A-corpus*. This time, examination at the level of a single text is more possible than it was for rhyme coefficients, such that here we exclude *Guillaume*, *Robert*, *Richard*, and *Jehan Paulus*, to which we

<sup>21</sup> *faire* is therefore the context lemma for these lemmata.

will return after discussion of the corpus' other texts. In this subset of the *A-corpus*, then, we observe thirty-two lemmata pairs which appear at the rhyme on greater than five occasions and which display a ratio of greater than 1. Five of these are also found for the *J-corpus*, namely *dire* twice (3.5 with *respit*, 2.3 with *martire*), *faire* twice (*debonnaire*, 2; *contraire*, 2), and the bivalent pair *mere-pere* (2.2 and 3). With the exception of *dire-respit*, we note that in these cases the ratios for the *A-corpus* are much lower than the associated ratios for the *J-corpus*, aligning with the assumption that texts in a single-author corpus are more likely to display consistent use of admixtures compared to a corpus of more diverse texts, written by authors with different preferences. This, along with the fact that the remaining twenty-seven lemmata pairs identified for the *A-corpus* (including a further six bivalent pairs) do not appear above our threshold in the *J-corpus*, strengthens the sense of cohesion for the *J-corpus*, and thus reinforces the shared-author hypothesis.

Turning our attention finally to *Guillaume*, *Robert*, *Richard*, and *Jehan Paulus*, these texts are too short on their own to permit a fine-grain analysis of the admixtures that appear within them through the calculation of ratios, but taking them individually, it is nevertheless notable that all four texts contain a number of the same admixture pairs as are also found in the *J-corpus*. *Guillaume*, *Robert*, and *Richard* share two pairs (*compagnie-mie* and *dire-respit*), while *Guillaume* and *Robert* share a further three (*faire-debonnaire*, *maniere-chiere*, and *mere-pere*), *Robert* and *Richard* a further two (*avant-enfant* and *grant-enfant*), and *Robert* and *Jehan Paulus* a further four (*compagnie-vie*, *folie-vie*, *maison-raison*, and *Marie-vie*). *Guillaume* displays three pairs (*dire-martire*, *dire-Jesus-Christ*, and *enfant-joiant*) which do not appear in the other texts, *Robert* contains seven (*clair-mere*, *dire-despit*, *frere-mere*, *mere-amer*, *plein-certain*, *Rome-somme*, and *vilain-certain*), *Richard* contains three (*avoir-remanoir*, *avoir-voloir*, and *faire-contraire*), and *Jehan Paulus* contains just one (*envie-vie*). There are thus only nine lemmata pairs of the thirty-four that meet our thresholds in the *J-corpus* which do not appear in any of these four texts.

In total, *Robert* displays the greatest alignment with the *J-corpus*, aligning for eighteen pairs compared to eight in *Guillaume*, seven in *Richard*, and five in *Jehan Paulus*.<sup>22</sup> The utility of this method is ultimately limited for these short texts, but it nevertheless still allows us to strengthen the growing body

**22** Explanation for the latter case in particular likely stems from the fact that *Jehan Paulus* is simply much shorter than *Guillaume*, *Robert*, and *Richard*.

of evidence that they ought to be attributed to Jehan, particularly for *Robert*. From a methodological perspective, meanwhile, there is an indication that this approach may be more usefully applicable for questions of authorship concerning longer poetic texts.

To conclude this chapter, we have seen that while historical claims in the scholarship as to the quality of Jehan's poetry are unhelpful, they nevertheless point to useful avenues through which his versification can be considered in a digital poetic analysis. Often, Jehan's tendencies come close to those observed in the comparable poems of the *A-corpus*, 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, itself supporting the shared-author hypothesis.

## Building the Case for the Shared-Author Hypothesis

Through the digital analysis of versificatory features in this chapter, including the rates of masculine and feminine rhyme use and the frequency of so-called “easy” rhymes, we have reinforced the internal coherence of the *J-corpus* when compared to the metrically and generically aligned texts of the *A-corpus*. We have also observed that, for these features, *Guillaume*, *Robert*, *Richard*, and *Jehan Paulus* sit squarely within the distribution of the *J-corpus*, reinforcing the single-author hypothesis. The examination of rhyme coefficients, while not suitable as a means of examining individual short texts, has nevertheless also added to the internal coherence of the *J-corpus*, and the distinctiveness of Jehan’s apparent preferences when selecting a word to appear at the rhyme. The most concrete evidence to support our new attributions, meanwhile, comes from the analysis of admixtures. *Robert* displays the strongest overlap with the *J-corpus* for such pairs, while *Guillaume*, *Richard*, and *Jehan Paulus* also display several admixtures which are hallmarks of the *J-corpus*, reinforcing the case for their inclusion, and suggesting that the *remanieur(s)* of *Richard* and *Jehan Paulus* left the ends of lines largely intact.

## Methodological Contribution

This chapter introduces a series of bespoke metrics, namely the concept of a *rhyme coefficient* and the analysis of admixtures, which have been developed to align with the specificities of poetry written in monorhyme alexandrine quatrains but have applicability in other contexts. By calculating how much more likely a word is to appear in rhyming position than elsewhere in the line, as well as by identifying pairs of words which are more likely to appear together, we can identify authorial habits that would remain obscure on a linear reading, and which nevertheless contribute to an individual versificatory fingerprint.



## SHARED VOCABULARY AND STRUCTURES

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 poems 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 fr. 24432.

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 plainement  
Les miracles qu'el fait a tous communement  
Qui la weilent servir de cuer parfaitement

(*Abesse*, vv. 9–12)

---

**ABSTRACT** Continuing our examination of repeated words and structures, this chapter considers what can be gleaned about Jehan's written style from an analysis of line-initial words and the structures used to mark the onset of discourse. In the former case, I consider the most frequent line-initial words, demonstrating that while individual texts display variability in the frequency with which these appear, the *J-corpus*, when taken as a whole, is statistically distinct from the *A-*, *F-*, and *R-corpora*. Following this, I propose a new typology for discourse onsets in the *J-corpus*, and perform an analysis which reveals further hallmarks of the *J-corpus* in the way that Jehan introduces direct speech, while also allowing us to observe differences in usage between the "long-text" sub-group of Jehan's *œuvre* and his shorter texts.

---

Here we may also repeat the similarly comparable stanzas found in both *Jehan Paulus* and *Chanoines*, which we identified above:

Si prions, touz et toutes, a la Vierge Marie,  
 Qui en ces flans porta le disgne fruit de vie,  
 Que tel oroyson soit de nous bouches parties,  
 Qu'avoir puissions des anges la sainte compaygnie.  
 (*Jehan Paulus*, vv. 367–70)

Si prions, touz et toutes, a la Vierge Marie,  
 Qui en ces flans porta le tres dous fruit de vie,  
 Qu'il nous doint si ouvrer en ceste mortel vie,  
 Par quoi aions des angres la sainte compaignie.  
 (*Chanoines*, vv. 409–12)

Morawski observes that many of the structures shared across Jehan's poems relate to shared elements of narrative,<sup>1</sup> 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,<sup>2</sup> while Mary appears as the one "qui le cors Dieu porta" in no fewer than five texts in the corpus.<sup>3</sup>

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 taking religious language as objective indicators of shared authorship since, by their very nature, devotional structures are likely to rely on clichés and commonplaces. As we have already argued in the case of *tous et toutes*, other structures can provide more convincing evidence of shared authorship, but this still relies on the assumption that the works in question were written by an author who had easy access to past output. Otherwise, as Munk Olsen suggests, we risk treating him like "un cerveau électronique dans lequel une impulsion de contenu donné déclenche automatiquement et infailliblement l'impulsion d'expression correspondante,"<sup>4</sup> in

1 For example, "Tost et isnelement entra en la chapelle" (*Povre chevalier*, v. 90; *Chevalier et escuier*, v. 107).

2 *Respon* (vv. 90, 176): "le roy de maesté"; *Pommes* (v. 306): "le roy de maisté"; *Abesse* (v. 131): "le roy de majesté."

3 *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*.

4 "an electronic brain in which an impulse of given content automatically and unfailingly triggers the corresponding impulse of expression."

other words a machine which, in consistent circumstances, will be incapable of varying its output.

As we argued above, even the stanzas shared between *Povre chevalier* and *Abesse*, and between *Jehan Paulus* and *Chanoines*, while encouraging as part of a broader case for shared authorship, are not enough on their own to definitively prove the shared-author hypothesis, since it is also plausible that the author of one text or another borrowed these relatively formulaic sections that they had read elsewhere. Indeed, given that the stanzas in both pairs are grammatically and narratively discrete, the inclusion could even have been made by a later scribe familiar with both texts. All of this is not to say that these shared structures are 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 in cases of shared narrative content, “ce sont toujours les mêmes expressions, formules et rimes qui reviennent en pareille occurrence.”<sup>5</sup> 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 fr. 24432, *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 one). This character turns their horse towards the chapel (phase two), and then enters the chapel and goes towards the image of Mary (phase three). 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 fr. 24432. Beyond the identical line already identified by Morawski and given above (p. 116n1), which appears during phase three, Kleist also notes that the stanzas containing the narrative phases two and three

---

**5** “it is always the same expressions, formulas, and rhymes that return in such instances.”

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).<sup>6</sup>

Some of Kleist’s evidence may be described as tentative, particularly the shared rhymes in phases two and three which contain “torna” and “chappelle,” both of which appear especially likely to appear in this narrative context. There is also the argument that narrative similarities such as these may have been the primary reason for the juxtaposition of both texts in fr. 24432, regardless of whether they share an author, rather than their proximity being an indicator of shared authorship. 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 such evidence, yet the combination of these observations with other similarities, not to mention the texts’ codicological context, certainly reinforces the possibility that there was an association 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 chapter, we will examine similar ways in which the shared-author hypothesis can be strengthened, or not, by the presence of shared structures across the *J-corpus*. In so doing, we will identify further evidence of structural coherence displayed by its texts beyond Morawski’s shared structures and Kleist’s points of narrative correspondence, allowing us to further progress the shared-author hypothesis.

## Line-Initial Words

Munk Olsen (1963b, 155) argues, counter to the suggestions of Morawski (1939, 1940) and Kleist (1973), 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 because of authorial preference. Instead, Munk Olsen argues, we should consider the use of

<sup>6</sup> “it is not just a matter of coincidental similarities, but of a stylistic principle that could be called the technique of set pieces.”

more semantically “empty” words, such as *et* and *quand*, particularly those at the beginnings of lines of verse, 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.”<sup>7</sup>

In doing so, Munk Olsen is aligned with developments in the methodologies which have come to be known as “stylometry,” pioneered by Mosteller and Wallace (1963, 275). Published in the same year that Munk Olsen released his first works on Jehan’s *œuvre*, and considering a dataset far removed from the present corpora, Mosteller and Wallace’s opening arguments are strikingly similar to Munk Olsen’s:

The filler words of the language such as *an*, *of*, and *upon*, and, more generally, articles, prepositions, and conjunctions provide fairly stable rates [in the work of a single author], whereas more meaningful words like *war*, *executive*, and *legislature* do not.

Mosteller and Wallace go on to show through statistical methods that “the function words of the language appear to be a fertile source of discriminators” (1963, 306). That is, as Munk Olsen suggested, it is those semantically “empty” words which best enable discrimination between potential authors for a given text.

Munk Olsen applies his method to some effect in his comparison of Jehan’s poems to the alexandrine *Alexis* included in our *A-corpus* (1963b, 159), using comparative counts of line-initial words to further his claim that the text cannot be considered part of Jehan’s *œuvre*. However, the 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). Further, 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 *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,<sup>8</sup> 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.”

---

**7** “it is more likely that they represent authorial ‘tics.’”

**8** *mais*, *car*, *lors*, *or*, *puis*, *adont*, *atant*, and *adoncques*, alongside the semantically “emptier” *et* and *quand*.

There is space, therefore, for a more data-driven approach to the same questions. And yet, as noted in our discussion on pp. 17–18, stylometry in its purest form falls outside of the scope of this study, the aim of which is to make use of explicitly bespoke methods, designed for a particular corpus of texts based on a humanistic understanding of those texts. With that in mind, an updated version of Munk Olsen’s approach, drawing from stylometry but using our understanding of the structure of alexandrine verse to look only at function words which come at the beginning of a line, promises to be more fruitful.

The focus here, as in our work on line-final forms in Chapter 6, 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 common, it also enables us to distinguish between uses of these same word forms which can serve as 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 *li*, *le*, *la*, *les*, etc. 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 34, the most striking feature is their variability.<sup>9</sup> 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*.

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 based on 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 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 to identify

---

**9** The decision to look specifically at the six most frequent lemmata is largely arbitrary, but examination of the most frequent lemmata is likely to yield the most reliable data.

Table 34. 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       |
|-----------------------------|-----|-------|-------|-------|-------|---------|
| <i>Respon</i>               | le  | il    | de    | que   | a     | et      |
| <i>Chanoines</i>            | le  | et    | a     | que   | quand | de      |
| <i>Pommes</i>               | le  | que   | quand | et    | a     | se      |
| <i>Romme</i>                | le  | de    | et    | que   | a     | il      |
| <i>Deux chevaliers</i>      | le  | que   | de    | et    | a     | lors    |
| <i>Enfant rosti</i>         | le  | et    | a     | lors  | son   | de      |
| <i>Povre chevalier</i>      | que | le    | de    | a     | car   | en      |
| <i>Chevalier et escuier</i> | et  | le    | je    | que   | quand | il      |
| <i>Narbonne</i>             | le  | a     | et    | lors  | quand | que     |
| <i>Chevalier hermite</i>    | le  | et    | que   | quand | de    | a       |
| <i>Cordouanier</i>          | le  | et    | de    | a     | en    | or      |
| <i>Juitel</i>               | le  | quand | et    | que   | a     | car     |
| <i>Enfant sauva mere</i>    | le  | et    | que   | a     | de    | quand   |
| <i>Eaue et vergier</i>      | le  | que   | et    | par   | de    | en      |
| <i>Pain</i>                 | le  | que   | et    | quand | de    | por     |
| <i>Mescreant</i>            | et  | le    | que   | il    | quand | si      |
| <i>Pecherresse</i>          | que | de    | le    | en    | cel   | quand   |
| <i>Merlin</i>               | le  | que   | et    | quand | de    | a       |
| <i>Florence</i>             | que | le    | mais  | a     | et    | quand   |
| <i>Anelés</i>               | que | le    | et    | mais  | de    | il      |
| <i>Buef</i>                 | le  | et    | que   | de    | a     | mais    |
| <i>Beguine</i>              | le  | et    | quand | que   | a     | de      |
| <i>Abesse</i>               | le  | que   | et    | quand | il    | seignor |
| <i>Sauveur</i>              | le  | quand | a     | de    | et    | que     |

whether there is indeed a unity to the frequency with which Jehan uses particular words at the beginning of each line.

Table 35 contains the total counts of the six words which appear most frequently within the *J-corpus*. These combined data appear to consolidate the trends observed in Table 34, 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 35. The six most common lemmata at the beginning of a line across the entire *J-corpus*.

| <i>le</i> | <i>que</i> | <i>et</i> | <i>de</i> | <i>quand</i> | <i>a</i> |
|-----------|------------|-----------|-----------|--------------|----------|
| 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 36. 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.

The first observation of note from these tests is that the variability that we observed anecdotally in Table 34 appears to be borne out by the data. Only seven texts do not display any statistically significant differences from the rest of the *J-corpus* for any of these lemmata, with nine differing for a single lemma, five texts for two lemmata, and two texts (*Chevalier et escuier* and *Sauveur*) for three lemmata. *Florence* differs from the usage observed for the rest of the *J-corpus* in no fewer than four 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 position 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 results from the fact that *le* is used even more frequently in these texts than we might expect, based on the rates observed across the rest of the *J-corpus*.

It would be tempting, at least initially, to assume that this variability provides evidence against the shared-author hypothesis. However, while we have already observed areas of significant variation within the *J-corpus*, we have also seen that this internal variation is often still less than that observed for our other reference corpora. Indeed, such variation also does not necessarily detract from the fact that the texts of the *J-corpus*, however internally variable, still form a cohesive unit that is collectively distinct from other corpora.

Table 36. Results of p-value tests comparing use of the six most frequent lemmata at the beginning of a line across the entire *J-corpus*.

| Text                        | <i>le</i> | <i>que</i> | <i>et</i> | <i>de</i> | <i>quand</i> | <i>a</i> |
|-----------------------------|-----------|------------|-----------|-----------|--------------|----------|
| <i>Respon</i>               |           |            |           |           |              |          |
| <i>Chanoines</i>            |           |            |           |           |              | PASS     |
| <i>Pommes</i>               | PASS      |            |           |           |              |          |
| <i>Romme</i>                | PASS      |            |           | PASS      |              |          |
| <i>Deux chevaliers</i>      |           |            |           |           |              |          |
| <i>Enfant rosti</i>         | PASS      | PASS       |           |           |              |          |
| <i>Povre chevalier</i>      |           |            | PASS      |           |              |          |
| <i>Chevalier et escuier</i> | PASS      |            | PASS      | PASS      |              |          |
| <i>Narbonne</i>             |           |            |           |           |              |          |
| <i>Chevalier hermite</i>    |           |            |           |           |              |          |
| <i>Cordouanier</i>          | PASS      | PASS       |           |           |              |          |
| <i>Juitel</i>               |           |            |           |           | PASS         |          |
| <i>Enfant sauva mere</i>    |           |            |           |           |              |          |
| <i>Eaue et vergier</i>      |           |            |           |           |              |          |
| <i>Pain</i>                 |           | PASS       |           |           |              |          |
| <i>Mescreant</i>            |           |            | PASS      |           |              |          |
| <i>Pecherresse</i>          | PASS      |            |           |           |              |          |
| <i>Merlin</i>               |           |            |           |           |              |          |
| <i>Florence</i>             | PASS      | PASS       | PASS      | PASS      |              |          |
| <i>Anelés</i>               | PASS      | PASS       |           |           |              |          |
| <i>Buef</i>                 | PASS      |            |           |           |              |          |
| <i>Beguine</i>              |           |            | PASS      |           | PASS         |          |
| <i>Abesse</i>               | PASS      |            |           |           |              |          |
| <i>Sauveur</i>              |           | PASS       | PASS      |           | PASS         |          |

By taking a more global approach to the *J-corpus*, we can build this sense of cohesion, by identifying ways in which Jehan's texts differ collectively from trends observed elsewhere. Here we compare the corpus to our three textually coherent reference corpora, namely the *A-corpus* of narrative poems in alexandrine monorhyme quatrains, the *F-corpus* of *fabliaux*, and the *R-corpus* of Rutebeuf's works. By comparing the *J-corpus* to these corpora, we thus compare it to poems of a comparable type to the *J-corpus*,

works of a consistent genre, and works by a single author. In doing so, we observe that, despite their surface-level variation, the texts of the *J-corpus* do, in fact, form a coherent corpus in how, collectively, they differ from these other corpora of text.

Table 37. The six most common lemmata at the beginning of a line across the *J*-, *A*-, *F*-, and *R-corpora*.

| <i>Corpus</i> | 1   | 2   | 3  | 4  | 5     | 6    |
|---------------|-----|-----|----|----|-------|------|
| <i>J</i>      | le  | que | et | de | quand | a    |
| <i>A</i>      | que | le  | et | de | a     | mais |
| <i>F</i>      | que | et  | le | a  | si    | de   |
| <i>R</i>      | que | et  | le | ne | de    | si   |

Table 37 displays the six most frequent lemmata in line-initial position for the *J*-, *A*-, *F*-, and *R-corpora*. By contrasting our reference corpora with the *J-corpus*, we notice in a superficial way that there is variation in which lemmata are the most frequent. While *quand* is the fifth most frequent line-initial lemma in the *J-corpus*, it drops from the list entirely for the other corpora, appearing in positions seven, twenty-three, and eleven for the *A*-, *F*- and *R-corpora* respectively. The use of *le*, meanwhile, is especially significant as it is the most frequent line-initial lemma in the *J-corpus* (by some margin, as we saw in Table 35), while it is only the second most frequent in the *A-corpus*, and third in the *F*-, and *R-corpora*.

We note that both these examples, while reinforcing the cohesion of the *J-corpus*, also align it with the *A-corpus* as the closest reference corpus, reinforcing the suggestion that at least some of this order may be determined by generic and versificatory requirements, even if authorial preference does play an additional role. The appearance of *mais* in sixth position for the *A-corpus* also reinforces this association, given that it also appears in seventh position in the *J-corpus*.

Considering the data from the other direction, we find several lemmata which appear with a high frequency in our reference corpora, but which do not appear as frequently in the *J-corpus*. Again, the association between the *J*- and *A-corpora* is strengthened by their comparable use of *et*, given that it is the third most frequent lemma in line-initial position in both, and yet it appears in second position in the *F*- and *R-corpora*. Similarly, *si* appears at a relatively high position for the *F*- and *R-corpora*, in positions five and six respectively, but only appears in positions fourteen and ten for the *J*- and *A-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 similarities of 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 by means of the specific rates at which these lemmata are used at the beginning of a line.

Table 38 contains the results of Welch Two Sample t-tests comparing the *J-corpus* to our reference corpora for each of the six most common line-initial lemmata within the *J-corpus*. Overall, and specifically for *le* and *quand*, the results reveal that many line-initial lemmata are used significantly differently in the *J-corpus*, when compared to the rates observed in our reference corpora.

Table 38. Average rates of the use of the six most common lemmata at the start of lines in the *J-corpus*, compared to the *A-*, *F-*, and *R-corpora*.

|              | <i>J-corpus</i> | <i>A-corpus</i> |         | <i>F-corpus</i> |         | <i>R-corpus</i> |         |
|--------------|-----------------|-----------------|---------|-----------------|---------|-----------------|---------|
| Lemma        | Rate            | Rate            | P-value | Rate            | P-value | Rate            | P-value |
| <i>le</i>    | 12.08%          | 8.64%           | 0.00001 | 6.01%           | 0.00000 | 5.07%           | 0.00000 |
| <i>que</i>   | 7.39%           | 9.36%           | 0.08833 | 12.15%          | 0.00000 | 13.85%          | 0.00000 |
| <i>et</i>    | 7.06%           | 7.79%           | 0.1778  | 10.01%          | 0.00040 | 9.28%           | 0.00397 |
| <i>de</i>    | 4.28%           | 3.94%           | 0.7611  | 3.02%           | 0.00672 | 3.93%           | 0.05749 |
| <i>quand</i> | 4.14%           | 3.16%           | 0.00912 | 1.40%           | 0.00000 | 2.09%           | 0.00000 |
| <i>a</i>     | 4.09%           | 3.54%           | 0.6083  | 3.07%           | 0.0124  | 2.73%           | 0.00003 |

In addition, we note that the *F-* and *R-corpora* are also distinct from the *J-corpus* for more lemmata (six and five respectively) than the *A-corpus*, which is only statistically distinct on two occasions. These data therefore reinforce the impression gained from the simple ordering of lemmata given in Table 34, that the *A-corpus* is most closely aligned with the *J-corpus*. This of course highlights the texts' narrative and versificatory similarities, but the presence of clear distinctions for *le* and *quand* also reinforces the fact that the *A-corpus* still differs from the *J-corpus* in a way that strengthens the shared-author hypothesis.

We are also able to draw broader linguistic conclusions from these data, namely that while uses of *le*, *que*, *et*, *quand*, and *a* vary significantly across all the corpora we have examined, the rate of use of *de* appears to be more

static, suggesting that the rates in this case may be 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 a valuable 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 reveal itself to be significantly distinct from that found in texts by other authors, across several metrics. Thus, Munk Olsen (1963b, 152) is correct, of course, to argue that an author is not straightforwardly “un cerveau électronique” with predictable outputs,<sup>10</sup> but as modern stylometry has proven in less bespoke ways, there are nevertheless macro-level trends which can unite a given corpus of texts through distinctiveness relative to some other corpus, in turn contributing to a sense that the texts share the same author.

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*, *Robert*, *Richard*, and *Jehan Paulus*, the texts of the *A-corpus* to which we have extended the shared-author hypothesis, line up with the trends observed for line-initial lemmata in the *J-corpus*. As we performed above for each of the texts of the *J-corpus*, Table 39 presents the results of p-value tests comparing these three poems to the *J-corpus*.

Table 39. Results of p-value tests comparing use of the six most frequent lemmata at the beginning of a line across the entire *J-corpus* to *Guillaume*, *Robert*, *Richard*, and *Jehan Paulus*.

| Text                | <i>le</i> | <i>que</i> | <i>et</i> | <i>de</i> | <i>quand</i> | <i>a</i> |
|---------------------|-----------|------------|-----------|-----------|--------------|----------|
| <i>Guillaume</i>    | PASS      | PASS       | PASS      |           |              |          |
| <i>Robert</i>       |           | PASS       | PASS      | PASS      | PASS         |          |
| <i>Richard</i>      | PASS      | PASS       | PASS      |           | PASS         |          |
| <i>Jehan Paulus</i> |           |            | PASS      |           | PASS         |          |

From these tests, we once again get the sense that for *Jehan Paulus* in particular, the text is sufficiently similar to the broader *J-corpus* to support collective attribution, given that it differs in a statistically significant way from the *A-corpus* on only two occasions, in the cases of *et* and *quand*, and is thus well within the range of variation observed across the *J-corpus*. This is of

10 “an electronic brain.”

course especially notable, given our suggestion that the text has at some point been *remanié*, and indicates that in this text, the later modifying poet made minimal changes to structures at the beginning of a line.

*Guillaume*, *Robert*, and *Richard*, however, appear somewhat more problematic. *Guillaume* displays statistically significant differences from the *J-corpus* for three lemmata (*le*, *que*, and *et*), and *Robert* (*que*, *et*, *de*, and *quand*) and *Richard* (*le*, *que*, *et*, and *quand*) do so for four lemmata each. If we compare these tests to those performed on the individual texts of the *J-corpus*, we therefore find that *Robert* and *Richard* sit alongside *Florence*, which is also distinct from the rest of the *J-corpus* in the case of four lemmata, and indeed overlaps on three of these in each case. One might therefore contemplate the exclusion of *Florence*, *Guillaume*, *Robert*, and *Richard* from Jehan's *œuvre*, but it is noteworthy that, as we have seen and unlike *Jehan Paulus* which sits more comfortably within the usual range, these four texts are amongst the longest which have been attributed to Jehan. Indeed, they constitute a group apart from the rest of his *œuvre*, alongside *Anelés* and *Buef*, forming what we have termed the "long-text" sub-group.

In part, this outcome is likely therefore a feature of the statistical methods used, since the t-test becomes more certain in its result the greater the sample size, and thus a longer text is more likely to trigger the threshold for statistical significance. However, there is more to this variation than simple variations in sample size. One final example will reinforce the cohesion of this grouping, allowing us to conclude that, just as Jehan's versificatory complexity increased when writing longer poems, resulting in lesser use of so-called "easy" rhymes, so too did the complexity of his line-initial structures, hence his longer texts are more likely to deviate from expected patterns seen elsewhere in the corpus. This, in turn, grants us a surer footing from which to attribute *Guillaume*, *Robert*, and an earlier version of *Richard* to Jehan.

The evidence for this claim comes from the lemma *mais*, which only appears in the top six lemmata in the three longest texts of the *J-corpus*: *Florence*, *Anelés*, and *Buef*, appearing at positions three, four, and six respectively. It also appears at a comparable position in *Guillaume*, *Robert* and *Richard* (two, three, and two respectively). That *mais* is used a great deal less in the shorter texts of the *J-corpus* therefore suggests that Jehan's word choices may have varied depending on features of the text he was writing, such as its intended length and thus its narrative complexity. When writing longer texts, Jehan appears to have preferred to use more contentful words at the beginning of lines, such as *mais*, perhaps to better engage an audience or simply to avoid significant repetition of standard structures. That this is true of all of the texts of the "long-text" group suggests that what matters more

than the texts' points of difference from the *J-corpus* as a whole is therefore the fact that they share notable similarities with other texts from the "long-text" group, to which we now add *Guillaume*, *Robert*, and a putative earlier version of *Richard*, the latter of which again does not appear to have been especially heavily *remanié* for this feature. We shall observe further evidence for the validity of this subdivision in the next section of this chapter.

The method pursued in this section has once more allowed us to affirm the shared-author hypothesis. It has done so while continuing to treat the texts in question in a humanistic way, acknowledging the undeniable truth 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 expanding this approach to consider other types of shared structures across the *J-corpus*. In doing so, we will gain a clearer idea of both the diversity and the unity of Jehan's texts.

## 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.<sup>11</sup> 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.

**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.

---

**11** Here I 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* 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.

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 respon: "Voulentiers le ferai." (*Chanoines*, 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 this passage from *Pain*, in which the direct discourse appears after an intervening line of indirect speech:

Li ennemis parole a l'ange d'une part  
 Por quoi Diex veult avoir en cestè ame part:  
 "Il m'est avis qu'il s'est aperceüs trop tart,  
 [...]" (*Pain*, vv. 113–15)

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 arguments 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 manuscript only through the content of the discourse.<sup>12</sup> These onsets of discourse may take place between one line and the next, or within the same line:

“Dame, vous estiez morte. Qui vous resuscita?”  
 “Eu non Dieu, dous amis, cele qui Dieu porta.” (Cordouanier, vv. 131–32)  
 “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:<sup>13</sup>

[...] 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:<sup>14</sup>

**Type A:** Discourse begins at the start of a new line, with the speaker indicated earlier.

**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.

**12** 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 fr. 24432 only rarely containing any punctuation to indicate the beginning of a unit of discourse. On 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 (2006a).

**13** When this occurs, both types of discourse onset are included in our counts.

**14** Types A and C combine to form a category of discourse onset which corresponds with what Cerquiglini (1981) and Marnette (2006a) 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.

**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, as a continuation of the same line or at the beginning of 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 40,

Table 40. Counts of each type of discourse onset for each text in the *J-corpus*.

| <b>Text</b>                 | <b>Type A</b> | <b>Type B</b> | <b>Type C</b> | <b>Type D</b> | <b>Total</b> |
|-----------------------------|---------------|---------------|---------------|---------------|--------------|
| <i>Respon</i>               | 6             | 11            | 10            | 6             | <b>33</b>    |
| <i>Chanoines</i>            | 11            | 8             | 12            | 1             | <b>32</b>    |
| <i>Pommes</i>               | 11            | 19            | 9             | 7             | <b>46</b>    |
| <i>Romme</i>                | 6             | 10            | 1             | 0             | <b>17</b>    |
| <i>Deux chevaliers</i>      | 5             | 16            | 7             | 1             | <b>29</b>    |
| <i>Enfant rosti</i>         | 3             | 11            | 4             | 0             | <b>18</b>    |
| <i>Povre chevalier</i>      | 4             | 1             | 7             | 12            | <b>24</b>    |
| <i>Chevalier et escuier</i> | 4             | 19            | 11            | 2             | <b>36</b>    |
| <i>Narbonne</i>             | 5             | 16            | 3             | 0             | <b>24</b>    |
| <i>Chevalier hermite</i>    | 6             | 9             | 9             | 1             | <b>25</b>    |
| <i>Cordouanier</i>          | 3             | 10            | 5             | 2             | <b>20</b>    |
| <i>Juitel</i>               | 9             | 2             | 5             | 4             | <b>20</b>    |
| <i>Enfant sauva mere</i>    | 4             | 9             | 13            | 0             | <b>26</b>    |
| <i>Eaue et vergier</i>      | 3             | 4             | 2             | 0             | <b>9</b>     |
| <i>Pain</i>                 | 4             | 5             | 7             | 0             | <b>16</b>    |
| <i>Mescreant</i>            | 0             | 11            | 12            | 0             | <b>23</b>    |
| <i>Pecherresse</i>          | 2             | 2             | 6             | 0             | <b>10</b>    |
| <i>Merlin</i>               | 3             | 3             | 12            | 20            | <b>38</b>    |
| <i>Florence</i>             | 13            | 9             | 50            | 0             | <b>72</b>    |
| <i>Anelés</i>               | 11            | 3             | 44            | 1             | <b>59</b>    |
| <i>Buef</i>                 | 13            | 23            | 17            | 1             | <b>54</b>    |
| <i>Beguine</i>              | 5             | 11            | 5             | 0             | <b>21</b>    |
| <i>Abesse</i>               | 3             | 23            | 4             | 3             | <b>33</b>    |
| <i>Sauveur</i>              | 15            | 32            | 5             | 2             | <b>54</b>    |
| <b>Total</b>                | <b>149</b>    | <b>267</b>    | <b>260</b>    | <b>63</b>     | <b>739</b>   |

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" (1963b, 160),<sup>15</sup> 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.

By considering the relative rates of use of each type per text, as represented in Figure 15, it becomes clear that there is a great deal of variation in the types used by individual texts, especially for Type D. Except for *Mescrèant*, which has no examples of Type A, all texts contain examples of Types A, B, and C. For Type D, ten of the twenty-four 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 *Anelès* 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. This impression of a negative correlation between Type B and Type C onsets is borne out by a Pearson's product-moment correlation test, which reveals a strong, negative correlation between the two onset types ( $-0.65328$ ,  $p = 0.00054$ ), such that in the *J-corpus* Types B and C can be seen as variants in alternation, with a high rate of use of one type leading to fewer possibilities for the other. This is a surprise, given the narrative difference between the proleptic Type C and the analeptic Type B,<sup>16</sup> but it 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 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.

A similarly strong, negative correlation is observed between Types B and D ( $-0.50540$ ,  $p = 0.01176$ ), such that a higher rate of one is associated

<sup>15</sup> "by far the most frequent."

<sup>16</sup> See p. 130n14.

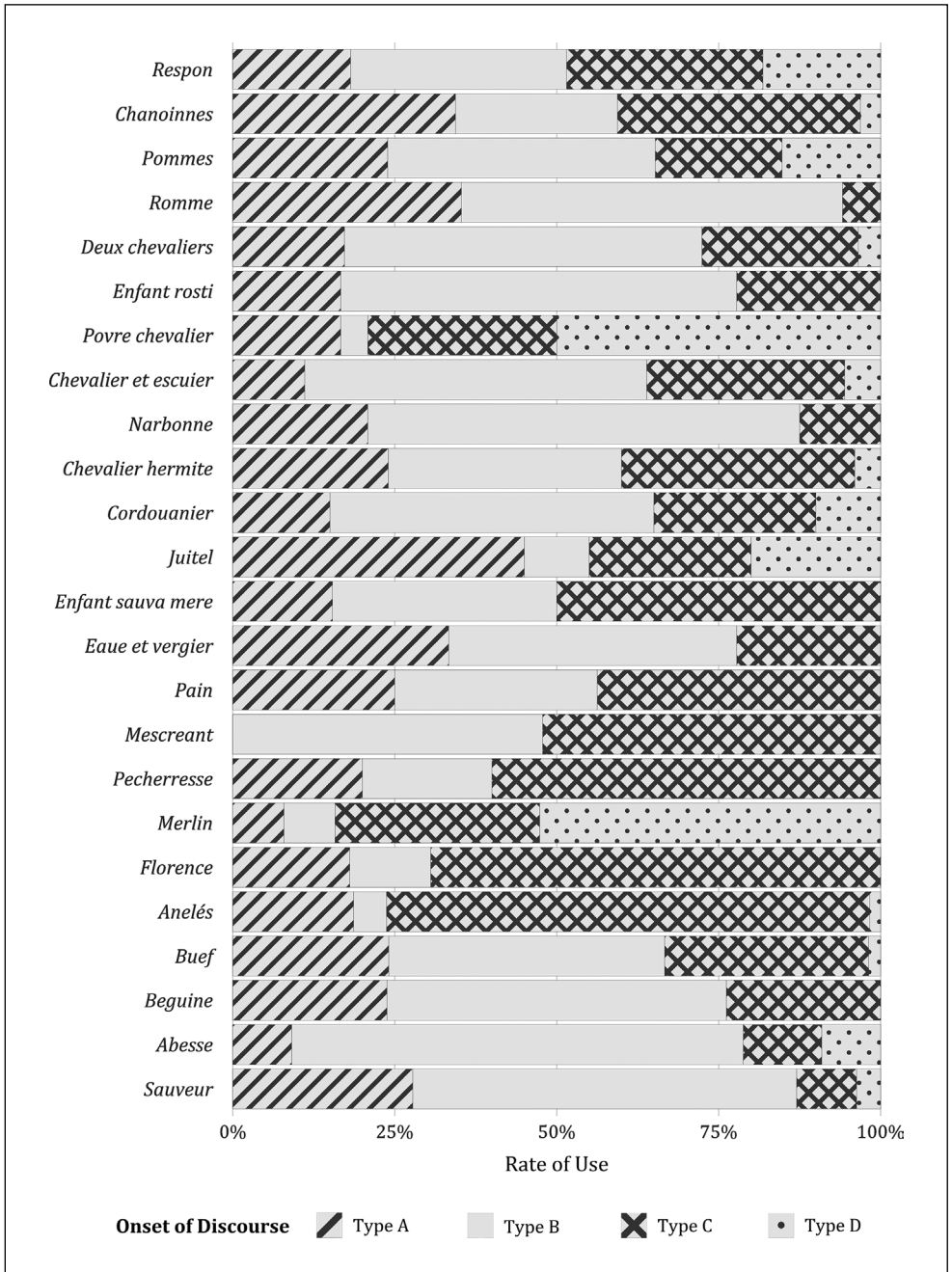


Figure 15. Rates of use of each type of discourse onset in the *J-corpus*.

with a lower rate of the other. This is again surprising, but can be explained through the fact that Types B and D in some sense seek to achieve the same goals: by placing the discourse as early as possible, it is granted a sense of immediacy. Then, once the discourse has begun, indication of the speaker can be given if it would be unclear in context (Type B), or the shift in speaker can be marked in more subtle ways (Type D). Looking again at our Type D example from *Cordouanier*, for instance, we see that “dous amis” functions as a marker of turn-taking, and indicates that the addressee is male (and therefore in this dialogue between a husband and wife, that the speaker must be female). In a more ambiguous case, the phrase could just as easily have been replaced with the trisyllabic “ce dist el” or similar, shifting the Type D onset to one of Type B.

“Dame, vous estiez morte. Qui vous resuscita?”

“Eu non Dieu, dous amis, cele qui Dieu porta.”

(*Cordouanier*, vv. 131–32)

A final correlation of significance is that of length. Given that *Florence* and *Anelés* are two of the longest in the *J-corpus*, and contain the greatest number of discourse onsets, their greater use of Type C onsets means that we are able to reinforce our hypothesis above that the longer texts within the *J-corpus* form a sub-group distinct from the rest of Jehan's *œuvre*.<sup>17</sup> A correlation test reveals that, across the *J-corpus*, there is a weak, positive correlation between text length and the rate of use of Type C onsets (0.48798,  $p = 0.01556$ ) such that, as a general rule, the longer a text is in the *J-corpus*, the higher its expected rate of Type C discourse onsets.<sup>18</sup> Notably, however, *Buef*, the other text within the *J-corpus* belonging to the “long-text” sub-group, while not completely out of alignment with these texts, nevertheless displays rates of use of Types B and C more in keeping with the rest of the corpus. The uses observed in *Florence* and *Anelés* are therefore explicable, and in line with a trend that exists across the whole corpus, but high use of Type C onsets is not a universal feature of the whole of this sub-group.

---

**17** 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; “by far the most frequent”) appears far more plausible, since we observe 255 instances of Type B, but only 166 instances of Type C.

**18** A correspondingly inverse correlation between length and Type B onsets appears plausible, but does not reach our threshold for statistical significance ( $p = 0.1492$ ). Neither Type A nor Type D onsets display a correlation with length at anything close to statistical significance.

Given these points of similarity between *Florence* and *Anelés*, we are reminded of Morawski's claim that "nous croyons que [*Florence* et *Anelés*] sont l'œuvre d'un même auteur, mais nous n'avons pas acquis la conviction qu'ils soient de Jean" (1939, 351–52).<sup>19</sup> The broader correlation between length and Type C onsets gives us reason to disagree with Morawski's claim, especially given the unifying features we have seen elsewhere. Our analysis here has nevertheless enabled us to encounter further stylistic similarities between these two texts, suggesting that if indeed they are the work of the same author, then they may have originated from a similar context of composition.

Comparing these data to those observed for the *A-corpus* in Figure 16,<sup>20</sup> we once again gain the sense that we have often seen for these corpora that many of the poems we are attributing to Jehan are broadly aligned with those of a similar poetic form and narrative style. In this case, this is true both in the distribution of different quote types, and in the very fact of their variability. As in the *J-corpus*, Types A, B, and C are prevalent in the *A-corpus*, with only *Déz* lacking instances of Type A, and *Confesseur* examples of Type B. All contain examples of Type C, while Type D discourse onsets are again much more variable, being absent in ten of the twenty-three texts of the *A-corpus* that contain direct discourse. *Guillaume*, *Robert*, *Richard*, and *Jehan Paulus*, meanwhile, again align with the rates we observed in the *J-corpus*, suggesting that even the apparent *remaniement* of the latter two texts had minimal impact on their use of discourse onsets. *Guillaume* and *Robert* display a comparable use of discourse onsets to *Anelés* and *Florence*, two texts of the "long-text" grouping alongside which *Robert* appears in fr. 24432, further strengthening the attribution of these texts.

We may therefore ask to what extent these two corpora do really differ from each other. A series of Welch Two Sample t-tests, which calculate the probability that the rates for each discourse onset in the *J-* and *A-corpora* are statistically distinct, here produce results (A:  $p = 0.226$ ; B:  $p = 0.1072$ ; C:  $p = 0.1228$ ; D:  $p = 0.2873$ ) that are often close to but all ultimately above our threshold for statistical significance. This remains true even if we exclude *Guillaume*, *Robert*, *Richard*, and *Jehan Paulus* from the *A-corpus*, or include them alongside the *J-corpus*. As a result, we can suggest that, overall,

<sup>19</sup> "we believe that [*Florence* and *Anelés*] are the work of the same author, but we have not reached the conviction that they are by Jean."

<sup>20</sup> Here we exclude *Thomas*, *Yves*, and *Dieu d'Amours*, which contain no examples of direct speech. In the case of the latter two texts, this is perhaps only because they are fragmentary.



Figure 16. Rates of use of each type of discourse onset in the texts of the *A-corpus* that contain instances of direct discourse.

Jehan's use of discourse onsets aligns with the expected patterns observed across other narrative poems from the same period, that were written in alexandrine monorhyme quatrains.

That is not to say, however, that discourse onsets are a feature totally removed from the question of authorial preference. As in the *J-corpus*, there is a strong, negative correlation in the *A-corpus* between the use of Type B and C onsets ( $-0.71602$ ,  $p = 0.00012$ ), once again suggesting that this is a feature determined by generic or versificatory expectations, but the same is not true of the correlation between Types B and D ( $p = 0.5824$ ). Again, this remains true even if we exclude *Guillaume*, *Robert*, *Richard*, and *Jehan Paulus* or move them to the *J-corpus*. The latter correlation therefore appears to be a feature only of the *J-corpus*, and thus while this is ultimately an indicator of variation, it is nevertheless evidence of an authorial preference, adding to our sense of cohesion for the *J-corpus*, and the sense of a distinction between the *J-* and *A-corpora* for this feature.

As a final direction of analysis, it is worthwhile examining discourse onsets in the *F-corpus* of *fabliaux*, to observe what this corpus may tell us about the extent to which the features we have observed thus far are limited to texts of this type, or shared with texts which are generically and metrically aligned, and yet collectively distinct from the *J-* and *A-corpora*. Whereas Welch two sample t-tests showed that there is no statistically significant distinction between the rates of any onset type when comparing the *J-* and *A-corpora*, when compared to the *F-corpus*, both display rates of use that are distinct in a statistically significant way for Types C (*J*:  $p = 0.00564$ ; *A*:  $p = 0.00013$ ) and D (*J*:  $p = 0.04029$ ; *A*:  $p = 0.00002$ ). We can therefore say with some certainty that the use of these types of discourse onset is determined by generic or versificatory features. Most likely, the contributing factor is the use of octosyllabic lines in the *fabliaux*, compared to the twelve-syllable alexandrine lines of the *J-* and *A-corpora*, which appears to result in a lower use of Type C onsets, which require greater space in the line, and an increase in Type D onsets, which do not. A statistically significant distinction also appears between the *A-* and *F-corpora* for Type B onsets ( $p = 0.01014$ ), which again aligns with the hypothesis that octosyllabic lines favour discourse onsets, like Types A and D, which allow for minimal material to appear in the same line as the discourse itself.<sup>21</sup>

**21** Indeed, even when Type B discourse onsets do appear in the *F-corpus*, the discourse often begins on one line, before the speaker is introduced in the following line. This does not occur in the *J-corpus*, and only very rarely in the *A-corpus*. Again, this is unlikely to have been the result of authorial choice, rather a function of the

Turning more broadly to the similarities in structure which we have observed for the *J-corpus* and beyond in this chapter, 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."<sup>22</sup> 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.

Overall, however, we have observed that there are 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, Robert, Richard*, and *Jehan Paulus*, allowing us to be more assured in the shared-author hypothesis.

---

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 marker, which is therefore pushed onto the following line.

**22** "an electronic brain in which an impulse of given content automatically and unfailingly triggers the corresponding impulse of expression."

## Building the Case for the Shared-Author Hypothesis

The statistical examination of line-initial lemmata and discourse onsets carried out in this chapter reinforces the internal coherence of the *J-corpus* by identifying the authorial “tics” initially discussed by Munk Olsen (1963b) which distinguish Jehan’s writing from that found in our other reference corpora. The “long-text” sub-group, including the newly attributed poems *Guillaume*, *Robert*, and *Richard*, is further unified by their shared approach to line-initial lemmata, and their comparable use of discourse onsets, in a way that differentiates them from Jehan’s shorter poems. Meanwhile, the fact that both *Richard* and *Jehan Paulus* remain statistically aligned with the *J-corpus* in both the use of line-initial words and discourse onsets, despite their likely *remaniement*, supports the status of both texts as echoes of earlier poems by Jehan, and suggests that the later *remanieur(s)* carried over units of direct discourse much as they appeared in Jehan’s work.

## Methodological Contribution

The approach presented in the first part of this chapter aligns closely with the core assumption of stylometric approaches to attribution, namely that focusing on semantically “empty” function words can have significant value when seeking to attribute authorship. However, by examining lemmata specifically at the beginnings of lines, as well as by examining the corpus-specific case of discourse onsets, I have moved beyond a purely stylometric approach to one which builds first and foremost from a humanistic understanding of the texts under consideration. My revised typology of discourse onsets, meanwhile, designed for the unique features of poems written in alexandrine monorhyme quatrains but applicable in other contexts too, offers an example of how one can quantify and compare an author’s use of specific narrative structures between and within corpora. In addition, this section also highlights the importance of accounting for corpus-internal variables when drawing conclusions on a given feature. For example, the differences in the use of discourse onsets observed between Jehan’s longer and shorter poems could easily be mistaken for evidence of multiple authors, were a unifying explanation for this variation not to be found in the correlation between text length and the use of discourse onsets.



## SOURCES

MUNK OLSEN (1978, xxi-cxx) offers a comprehensive examination of the possible sources of the texts in the *J-corpus*, which I summarize here, before turning as our final source of evidence to an examination of the contribution that can be made to this question by digital textual analysis.

For several 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 (Munk Olsen 1978, lxxxv),<sup>1</sup> while *Chanoines*, *Juitel*, and *Merlin* are all based on tales found within the *Vie des peres*.<sup>2</sup> *Respon*, *Romme*, *Chevalier et escuier*, *Pecherresse*, 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 eleven tales appear to be original combinations of elements found elsewhere, existing in an

---

<sup>1</sup> For an edition, see van Hamel (1885).

<sup>2</sup> 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 *Chanoines*, see Chaurand (1971), for *Juitel*, see Wolter (1879), and for *Merlin*, see Méon (1823, vol. 2).

---

**Abstract** The final chapter of this study examines Jehan's interactions with his source material, considering claims of both oral and written sources. Using digital textual analysis, I carry out a comprehensive examination of these references as they appear in the *J-corpus*, and compare these to claims found in the *A*-, *B*-, and *F*-*corpora*. Statistical analysis reveals that the *J-corpus* is notable for its high rate of references to written sources, a feature also observed in *Guillaume* and *Robert*, which further reinforces the shared-author hypothesis. I also argue against the blanket rejection by previous scholarship of Jehan's claims to have consulted such written sources, suggesting that while his references to writing are generally cursory and intended primarily to increase the authority of his material, this is not universally the case. Instead, it is possible in some cases that Jehan's inspiration was more complex, and that he drew from both oral and manuscript sources, sometimes even for the same story.

---

established narrative tradition, with more or less significant modifications, namely *Deux chevaliers*, *Enfant rosti*, *Povre chevalier*, *Narbonne*, *Chevalier hermite*, *Enfant sauva mere*, the first tale in *Eaue et vergier*, *Pain*, *Mescreant*, *Beguine*, and *Sauveur*. In addition to this, *Pommes*, *Florence*, and *Anelés* resemble pre-existing *vitae*, with modifications and additions that do not appear elsewhere,<sup>3</sup> forming a group of reworkings of existing textual traditions focused on a single pious individual. To this group, we may also add our additional texts *Guillaume*, *Robert*, *Richard*, and *Jehan Paulus*. *Guillaume* is a reworking of a longer epic, which has been attributed with debatable validity to Chrétien de Troyes.<sup>4</sup> *Robert*, meanwhile, is apparently based on an earlier *roman*,<sup>5</sup> and Conlon (1977, 25) identifies various earlier parallels for the story told in *Richard*, and suggests that there may well have been “an earlier chanson or roman devoted to [Richard’s] adventures,” even though no such text survives. *Jehan Paulus*, meanwhile, is apparently connected to a thirteenth-century version of the same story, written in octosyllables.<sup>6</sup> 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 short summaries of didactic stories apparently intended as an *aide-mémoire* for sermon writers.<sup>7</sup> Only *Cordouanier* and *Buef* have no known analogues.

Overall, then, Jehan operated within his existing textual context, and his originality comes not in the creation of new stories, but in his *remanie-ment* of existing ones. He abridges longer tales and expands on shorter ones, all with the intention of promoting his brand of narrative didacticism. The poet himself acknowledges having drawn inspiration from sources which he encountered, but in ways which offer little direct evidence, and through claims which scholarship has refuted. Kleist (1973, 53) and Munk Olsen

---

**3** *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. 153n28, Wallensköld (1907), and Beaussart (2000). *Anelés* closely resembles *La Fille du comte de Pontieu*, edited by Brunel (1926). The “ring in fish” motif crucial to *Anelés* (Tubach 1969, §315) also appears within *Leu* and *Grégoire* in the *A-corpus*, in the latter case with a key rather than a ring.

**4** 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, and for an edition, see Short (2021).

**5** For an edition, see Gaucher (2006).

**6** For an edition, see Allyn Williams and Allen (1935, 82–140).

**7** On this, see our examination of *exempla* on pp. 46–53.

(1978, cxix) note the presence of a number of claims regarding the written sources of the *J-corpus*, with Kleist arguing that “[z]war 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.”<sup>8</sup> Similarly, Munk Olsen argues that any suggestion within Jehan’s texts that his primary sources were written are little more than “des clichés commodes,”<sup>9</sup> claiming that Jehan’s sources for his poems were in fact primarily oral, but that this was a reality which the author attempted to conceal.

## Claimed Oral Sources

Approaching this question of oral versus written sources through a digital lens can 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 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 much rarer. There are, in fact, only two instances in the *J-corpus* in which Jehan suggests that material for his poems came to him by oral means. The first of these is within *Anelés*, in which Jehan claims that “moult l’ay oï conter / Que ja ne perira ce que Diex veult sauver” (vv. 555–56). 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 encountered the texts which are believed to have been the sources for his own poetry.

The only example of such a claim about having heard the details of an entire text 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 it as a genuine indication of the poem’s source material. Even if this is not the case, then it does at least foreground the supposedly religious origins of the tale, while signposting Jehan’s awareness of the position of his poetry in an

---

8 “[a]lthough Jehan often refers to a written source, such references are formulaic and simply indicate that the author believes himself to be in accordance with written tradition.”

9 “convenient clichés.”

ongoing tradition of oral reception and performance, which is aligned with, but not identical to, the use of *exempla* in sermons. This being the only reference anywhere in the *J-corpus* to suggest that the primary source for the text was oral, the evidence does therefore appear to support Kleist and Munk Olsen's claims that Jehan prefers not to cite oral sources, even if this may indeed have been his method.

Examination of the *A-corpus*, where we only find two such examples, reveals that Jehan is not alone in avoiding such references. The first, in *Sebastien* (v. 249), is another straightforward reference to oral wisdom, akin to that found in *Anelés*, where the narrator states that "Vous oez un proverbe souvent, qui tout est voir," before citing it directly. A more substantial case comes in *Thomas*, with the narrator this time situating the story within a broader process of hearsay:

La novele en fu tost par tute la contree,  
Mais l'evesque ne pout creire la renumee,  
Desque celui li out s'aventure mustree.  
E il la nus conta, cum cil li ot cuntee.

(*Thomas*, vv. 73–76)

It is clear, then, that sources in the *J-corpus* are rarely given as oral, but the same is also true of the *A-corpus*, suggesting that this is a broader generic feature of the poems in both corpora.<sup>10</sup> This is not, therefore, as distinguishing a feature of the *J-corpus* as Kleist and Munk Olsen have suggested.

## Claimed Written Sources

### Reading

While references to texts experienced through oral performance are limited in the *J-corpus*, Jehan makes a good deal more claims to have gathered some of his material from written media. For example, across the poems of the *J-corpus*, we find three variants of *trouver lisant* to suggest the written nature of his sources, which are given in Table 41.

All three examples appear somewhat out of the blue, in sections of the text that set 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 those we observed to oral ones, they are

---

**10** There are twelve direct references to oral sources in the *F-corpus*, which is again lower than might otherwise be assumed, but these still appear at a higher rate than in the *J-* and *A-corpora*.

Table 41. Uses of the phrase *trouver lisant* in the *J-corpus*.

| Text                     | Line | LH                            | KW              |
|--------------------------|------|-------------------------------|-----------------|
| <i>Enfant rosti</i>      | 45   | Huit jors devant Noël, si com | trouvons lisant |
| <i>Narbonne</i>          | 18   | A l'escole fu mis, si com le  | truis lisant    |
| <i>Enfant sauva mere</i> | 29   | Il fu .i. riches hons, si com | trouvons lisant |

not sufficient on their own to allow us confidently to refute Munk Olsen's hypothesis that all such references are simply "des clichés commodes,"<sup>11</sup> included to add a sense of authority to Jehan's words, rather than a specific reference to his sources. Indeed, we 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, preceded by *si com* and with *lisant* at the end of the line, suggesting possible use simply as means of filling half 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* (v. 128), this time without a form of *trouver*, but again with a variant of *si com*, in "si comme nous lisons." 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, were used by Jehan predominantly as a way of adding authority to his poem while also filling an entire hemistich and providing a common rhyme.

Table 42. Uses of the phrase *trouver lisant* in the *A-corpus*.

| Text           | Line | LH                        | KW                 |
|----------------|------|---------------------------|--------------------|
| <i>Thibaut</i> | 127  | Helyas fu premiers, co    | trovons nos lisant |
| <i>Thibaut</i> | 417  | De caus de fors aucun, ce | trovons nos lisant |

Jehan is not unique in his use of *trouver lisant*, but it is notable that the two other examples that appear in the *A-corpus*, both within *Thibaut*, differ in their exact structure. In both examples, given in Table 42, we observe the presence of the pronoun *ce* before the verb, and the interposition of *nos* before *lisant*. These examples therefore do not encourage the attribution of

11 "convenient clichés."

*Thibaut* to *Jehan*, rather they reinforce that it is the entirety of the phrase *si com trovons lisant* which was preferred by *Jehan*, thus adding coherence to the *J-corpus* and reinforcing the attribution of *Guillaume*.

Finally, we may also argue that the verb *trouver* alone is a term more generally associated with reading in *Jehan's* texts, through association with its use in *trouver lisant*. A number of examples of this use 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 reading, which suggests again the use of such references as a means of asserting textual authority, and adds further unity to *Jehan's* texts.<sup>12</sup>

## Writing

Looking at references to *Jehan's* sources having been written down that appear in the *J-corpus*, as well as in *Guillaume* and *Robert, Richard, and Jehan Paulus*, we find further examples which create a body of evidence against the absolute nature of Kleist and Munk Olsen's claims regarding the oral nature of his sources.

As an initial point of comparison, we note that of the 101 texts in the *F-corpus* of *fabliaux*, only nine contain references to a written source for the text (8.9%). By contrast, across the *J-corpus*, *Guillaume*, *Robert, Richard*, and *Jehan Paulus*, which together make twenty-eight texts, there are some eleven texts which contain references to written sources (39.3%,  $p = 0.00634$ ). These references are made across seventeen individual examples, included in Table 43, 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.<sup>13</sup> The statistical significance of the rate at which such references appear points to a clear distinction between the *J-corpus* and the *F-corpus*, reinforcing the coherence of the *J-corpus*, while again reinforcing the strength of our attribution of *Guillaume* and *Robert* in particular to *Jehan*.<sup>14</sup>

<sup>12</sup> In the *A-corpus*, *Thibaut* also contains the line “Hissi lo trovons nos en la divinité” (v. 642).

<sup>13</sup> This is in addition to five references to writing within the narrative, and two references to Holy Scripture without discussion of how its content has influenced the narrative.

<sup>14</sup> No non-narrative references to writing appear within *Richard* or *Jehan Paulus*. If such references ever existed in *Jehan's* putative earlier versions, then these were perhaps removed by the later *remanieur(s)*.

Table 43. References to written sources in the texts of the *J-corpus*, as well as *Guillaume* and *Robert*.

| Text                         | Line | LH                                          | KW         | RH                                    |
|------------------------------|------|---------------------------------------------|------------|---------------------------------------|
| <i>Chanoines</i>             | 1    | L'                                          | escripture | tesmoingne et nous fait asavoir       |
| <i>Romme</i>                 | 81   | Elle fist tant de biens si com dit li       | escriis    |                                       |
| <i>Romme</i>                 | 103  | Et de ces .ii. pechiez si com dit li        | escriis    |                                       |
| <i>Deux chevaliers</i>       | 13   | Il nous est tesmoignié en la sainte         | Esriture   |                                       |
| <i>Eaue et vergier</i>       | 9    | La divine                                   | Esriture   | nous en fait mencion                  |
| <i>Pain</i>                  | 13   | Ainsinc com nous trouvons en la sainte      | Esriture   |                                       |
| <i>Mescreant</i>             | 21   | En la <i>Vie des peres</i>                  | escript    | en parchemin                          |
| <i>Florence</i>              | 20   | Que l'                                      | escripture | dit de ce ne doutez mie               |
| <i>Buef</i>                  | 87   | Grant penitence en firent si com dit li     | escriis    |                                       |
| <i>Sauveur</i> <sup>15</sup> | 4    | Ainsi est il                                | escript    | en la divinité.                       |
| <i>Sauveur</i>               | 11   | Or advent une fois, si come dit l'          | escriis    |                                       |
| <i>Guillaume</i>             | 7    | L'                                          | escripture | tesmoingne que qui vault hault monter |
| <i>Guillaume</i>             | 9    | Et c'est                                    | escript    | c'un riche qui ne veult dire rien     |
| <i>Guillaume</i>             | 29   | Le philosophe dist—c'est                    | escript    | en maint lieux—                       |
| <i>Guillaume</i>             | 141  | Que l'                                      | escripture | dit c'om doit pour Dieu laisser       |
| <i>Guillaume</i>             | 582  | Que je croi qu'il n'est clers qui le peüst  | escrire    |                                       |
| <i>Robert</i>                | 96   | Ainssi comme l'en trueve en plusieurs lieux | escriit    |                                       |

<sup>15</sup> 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 these structures contain. *Guillaume* shares “L’escripture tesmoigne” with *Chanoines* 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 aligns in metre such that the differences between the two structures could simply be scribal. 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).<sup>16</sup> *Guillaume* and *Robert* also share a similar structure, with “c’est escript en maint lieus” in *Guillaume* resembling “en pluseus lieus escrist” in *Robert*.

Notably, none of the examples in Table 43 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 thirteen, 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. First, as has already been observed for *trouver lisant*, eight of the seventeen examples above 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* (four examples), functions as another hemistich-length phrase which are simply tagged on to a line to provide an appropriate rhyme, while also adding a sense of textual authority. The structure is therefore once again one which contributes to the coherence of Jehan’s *œuvre*, but does little to reinforce the written nature of his sources.

A similar argument can be made for the second structure which appears at the end of a line, *en la sainte Escripiture* (two examples), but in this case the question of function is a little less straightforward. This phrase is not used simply to fill a line, since both examples appear as part of wider phrases which emphasize the reference to a narrative source. The possible function of these 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* (two

---

**16** I make no claim as to whether these phrases are oblique self-references by Jehan, not wishing to engage in the largely semantic debate around whether Jehan was a *clerc* (as Munk Olsen [1978, xi–xii] claims) or a *jongleur* (as Faral [1910, 176] posits).

examples, both in the repeating structures discussed above) or *divine* (two). Interestingly, the two references to Holy Scripture elsewhere in the *J-corpus* also make use of the descriptor *divine*,<sup>17</sup> with one such example appearing within *Pain*, which itself contains a description of Jehan's supposed source for the poem as *la sainte Escripature*. In the examples involving *divine*, the references are to general moralizing claims rather than to a narrative source. 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, as with *si com dit li escriis*, act as little more than markers of this authority, rather than invitations for an audience to go and seek out the sources themselves.

This also aligns with the four examples in Table 43 found towards the beginning of *Guillaume*, in which the narrator refers to written sources to reinforce the 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 once again 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.

This is not necessarily true for all 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 in all cases 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.<sup>18</sup> Jehan claims that the source for his version

**17** *Enfant sauva Mere* (v. 14) and *Pain* (v. 7).

**18** 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).

of this text is the *Vie des peres*,<sup>19</sup> itself apparently “escript en parchemin.” Again, this is clearly an appeal to authority, but it would seem premature to so easily dismiss this example, as Munk Olsen (1978, cxix) does, as a reference which does not “correspondre à [une] réalité.”<sup>20</sup>

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’s fourth-century *De schismate donatistarum*, in which heretics are punished for having fed the Host to their dogs.<sup>21</sup> 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.

Although no text survives which could offer such a link, it is 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 the vernacular, 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 no such version ever existed.

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

---

**19** 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 these references are to the same sequence of texts as the *Vie des peres* referenced in *Mescreant* (be they the texts known today as the *Vie des peres* or some other sequence).

**20** “correspond to reality.”

**21** Book 2, chap. 19: *Eucharistia a Donatistis canibus data, ampulla chrismatis in praeceptis projecta aliaque eorum sacrilegia*, PL 11 (Migne, 1845), cols. 972–74.

on fifteen occasions, with only three referring to the sources of a text.<sup>22</sup> Particularly comparable to Jehan's use is the structure found in the prologue to the *Lai de Tyolet*,<sup>23</sup> 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

(*Lai de Tyolet*, vv. 29–32)

In this case, the narrator claims an intervening role played by the Bretons, who:

[...] en firent lais plusors,  
si con dient nos ancessors

(*Lai de Tyolet*, 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. Despite the lack of any reference to intervening oral sources in *Mescreant*, unlike in *Tyolet*, it is possible that Jehan's claim in *Mescreant* is similar to that of the *lai*, 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 oral. Our conclusion for the text's source must therefore be tentative, but it would nevertheless be unwise to discount the possibility, as Munk Olsen does, that Jehan's primary source was written, or that he had at least encountered a written version of the story. We ought also, of course, not to assume that Jehan would only have encountered a given tale once, and in one medium: he could in fact have come across the same tales on several occasions, and in different media.

Taylor (2000, 740) offers some useful context here, namely that preachers using *exempla* of the kind which appear to have inspired some of Jehan's poems "usually informed the audience of the source of their examples," and used phrases like *j'ai vu* to claim that the tale arose out of personal

<sup>22</sup> Reference to parchment as a transmission medium for the tale being told appears in the *A-corpus* within *Confesseur*, in the line "Nul ne pourroit escrire en parchemin" (v. 104). The reference is reminiscent of that found in *Guillaume* (v. 582), but this time the reason for being unable to express the story in writing is because of the presence of great joy, not sorrow. It also offers little about the possible source of the poem in question.

<sup>23</sup> For an edition of this text, see Walter et al. (2018).

experience, *j'ai entendu* to suggest second-hand knowledge, and *lit dans* to indicate written sources. Jehan's frequent references to written sources, not least the specific example in *Mescreant*, therefore sit within a broader tradition of such citation within didactic narrative, which we would be unwise to dismiss out of hand. Similarly, Brémond, Le Goff, and Schmitt (1982, 63–64) note that as the fourteenth century progressed, a great deal of the reception of *exemplum*-type narratives was through the written, rather than the aural medium, in the form of moralized tales intended to be read, such as Nicolas Bozon's *Contes moralisés* (ca. 1320),<sup>24</sup> which take a tale akin to an *exemplum* as their starting point, but offer additional moral explanation surrounding it, as Jehan does. While Jehan would be early in this process of change in the *exemplum* from an oral to a written tradition, if indeed he did write at the turn of the fourteenth century, this possibility suggests that for the case of *Mescreant*, it is at least feasible that Jehan encountered the tale which inspired it through a written text which has now been lost.

We therefore must at least 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, he compares them in content to the chronicles of Saint-Denis:<sup>25</sup>

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)

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*,<sup>26</sup> 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

**24** For an edition, see Toulmin Smith and Meyer (1889). Indeed, we may choose to see *conte moralisé* as the most appropriate generic descriptor for Jehan's texts.

**25** 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.

**26** For an edition, see Weiske (1991–1992).

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),<sup>27</sup> such that the text may be seen to be closer to the thirteenth-century *Chanson de Florence de Romme* (2000, 8).<sup>28</sup>

A valuable point of comparison here from the *A-corpus* comes in the form of an extraordinarily detailed and yet seemingly wholly fictitious account of the translator’s art, found in the prologue of *Basile*:

L’istore, dont je voel or comencier mes dis,  
 Conmenche *approbate consuetudinis*.  
 Et qui ne m’en voet croire, voit lire les escriis,  
 Il trouvera pour voir tout ce que je vous dis.

L’istore doit bien yestre devant clers rechitee  
 Et devant laie gent ne doit yestre celee,  
 Car elle est de latin en francois translatee  
 Sans mettre et sans oster fors tant qu’elle est rimee.

Jou ay au translater mise grant estudie  
 Et, pour le mieus rimer, villiet mainte nuitie  
 Que de toutes gens fuist plus volentiers oïe.  
 Je voel or commenchier. Dieus m’en soit en aïe! (Basile, vv. 9–16)

Here, the poet-narrator goes much further than Jehan ever does in laying out his apparent sources and poetic method. He gives an *incipit* line for the text he claims to have translated, which is identifiable as a reference to a sermon composed by Fulbert of Chartres on the Nativity of the Virgin Mary. The sermon contains a very cursory reference to the life of Saint Basil, but as Denomy (1956, 105) notes, it is very unlikely to be the direct source of the narrative presented here, and *Basile* is certainly not a direct translation of it, as the poet-narrator claims. Like Jehan’s apparent reference to the *Gesta Romanorum*, then, the poet-narrator here is attempting to burnish his credentials by claiming that his poem is further up the chain of sources than can be considered plausible, as a means of reinforcing the strength of his didactic claims.<sup>29</sup>

**27** “is not exclusively inspired by the *Gesta*.”

**28** 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, 2:1–128).

**29** The same tendency to exaggerate one’s proximity to the earliest sources of a

A second reference to the chronicles in *Florence* reinforces this sense that Jehan seeks to show off his literary awareness and shore up the claim that his version of the tale is superior to others. When Florence's abusers arrive at Beaurepaire to meet their judgement, Jehan asserts:

Douce gent es croniques de Romme sont trouuees  
 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 recorderees  
 (*Florence*, vv. 693–96)

On this occasion, Jehan criticizes 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 moment in the tale. 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 collection,<sup>30</sup> 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, 5n5).

Whether or not we believe Jehan's claim to have encountered this "rommans," we may interpret his reference as an attempt to see off literary competitors telling similar stories. Indeed, it is also conceivable 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 own 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

---

tale appears in the work of compilers of *exempla*, who "se réclament de l'autorité et du prestige en revendiquant des sources anciennes [lorsqu']en vérité, beaucoup d'*exempla* sont issues de compilations intermédiaires" (Rehr and Polo de Beaulieu 2021, §11; "claim authority and prestige by citing ancient sources [when] in truth, many examples come from intermediate compilations").

**30** On which see Wallensköld (1907–1909).

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 trouuees,” 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.

## 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, and that he acknowledged his position in an extended textual tradition, allows us also to 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,”<sup>31</sup> and which Cholette (2022, 102) suggests is evidence of the poem presenting its sources as exclusively oral. A little later, however, we find an instance of the structure “si con dit li escriz” (v. 87), as given above in Table 43.

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),<sup>32</sup> 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*, has cast doubt on the strength of Munk Olsen’s claim that there is no way Jehan could ever have encountered his sources through a written 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*, *Robert*, and perhaps also *Richard*, Jehan’s tendency in the case of his longer texts is far more often to shorten even lengthier tales, rather than to lengthen shorter ones, and thus a brief sermonic *exemplum* is unlikely to have been the sole source for *Buef*, itself the second longest text in the *J-corpus*.

31 “certainly an exact description of his working method.”

32 “convenient clichés that should not be taken seriously.”

Overall, then, a more thorough analysis of the corpus than has previously been possible, enabled by digital methods, has allowed us to discover 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* in particular,<sup>33</sup> 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. While references to the written sources of a text are far from unique to Jehan, we are nevertheless thus able to add this evidence to the overwhelming sense of unity which we have observed in Jehan's texts up to this point, building the case for the shared-author hypothesis.

---

**33** We note again here that the lack of references to sources in *Richard* and *Jehan Paulus* could indicate that, unlike many of Jehan's earlier features which these texts appear to retain, references to their sources may well have been removed by the later *remanieur(s)*, wishing to present the reworked versions as innovative.

## Building the Case for the Shared-Author Hypothesis

The final unifying thread of this study comes in the form of Jehan's appeals to textual authority when presenting the sources of his tales. The texts we have attributed to Jehan, comprising the *J-corpus*, *Guillaume*, *Robert*, *Richard*, and *Jehan Paulus*, collectively display a much higher rate of claimed written sources than both the *A-* and *F-corpora*, and specific formulaic phrases used to enhance textual authority in *Guillaume* and *Robert* also serve to link these texts directly to others in Jehan's *œuvre*. The absence of references to writing in *Richard* and *Jehan Paulus*, while not evidence against attribution, could nevertheless suggest that such references, if they had existed in Jehan's putative earlier versions, were subsequently removed by the texts' later *remanieur(s)*.

## Methodological Contribution

In this final chapter, I have demonstrated how digital methods can assist in quantifying rhetorical strategies and appeals to textual authority across a broad corpus. By conducting statistical analysis on the relative frequency of references to written sources between corpora, I have displayed how the very act of claiming textual authority can be a more significant to questions of authorship than determining whether or not those claims are based on fact. Once again, by examining specific examples revealed through digital textual analysis, I have shown how data-driven approaches to text are at their most effective when combined with close textual engagement, allowing for a more holistic and falsifiable approach than would be possible using either method in isolation.



## CONCLUSION

AS WE HAVE SEEN throughout this study, there is no silver bullet within either digital or traditional methods that can prove beyond doubt that Jehan wrote these and only these poems. Our best chance of accurately attributing texts to him comes instead from a holistic, cumulative 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.

[A]n isolated criterion proves nothing, or rather can prove anything, which is even worse. The only adequate method is to establish a convergence of similarities between the [...] works.

This is exactly what I have attempted in this study, in which digital and statistical methods of analysis unavailable to previous generations of scholarship have allowed us to widen the variety of evidence brought to bear on the shared-author hypothesis. In any era, the presence or absence of a particular feature in a text cannot, in any concrete, 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, reworking, and reuse. Yet, by taking a global, holistic, and humanistic view of a variety of different strands of analysis, we gain a much clearer picture than any one strand alone might suggest.

Through the examination carried out in this study, 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 certain types of discourse onset. 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. Our approach has also facilitated the attribution to the same author of a further two texts in their current form, namely *Guillaume* and *Robert*, as well as the putative earlier forms of *Richard* and *Jehan Paulus*. These attributions have been made not based on any one point of similarity, but rather following the holistic examination of a large amount of diverse evidence.

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 within the work of a given author cannot be determined only in isolation, and can instead only be reached after a careful reading of a text and close engagement with its context. Texts are not just bags of words, and any approach that treats them as such must go hand in hand with more interpretive methods.

Beyond the question of Jehan's authorship, we have seen that the methods of analysis used in this study are broadly applicable to a wide array of Medieval French texts. Through our examination of the *F-* 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 study 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.<sup>1</sup> 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 Medieval French, such as that of the *Lai d'Aristote*,<sup>2</sup>

---

**1** 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 shed further light on the matter.

**2** 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

as well as several of the more peripheral texts attributed to Chrétien de Troyes.<sup>3</sup>

Analysis of these cohesive 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 more standard feature within medieval French. 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 narrative poems written in alexandrine monorhyme quatrains, the *A-corpus* in particular has revealed that form and genre play a 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 study let us end, as we began, with an examination of the question of why, if at all, one would attempt 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 poet named Jehan de Saint-Quentin who composed 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 our analysis, 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 so as a means of determining shared authorship, but these features may also be fruitfully considered from a more literary perspective. Such an approach allows us to encounter Jehan's *œuvre* as a network of interconnected texts, each of

---

Olsen (1963a, 1963b, 1978) for Jehan's texts, but which does not venture into the methods presented in this study.

**3** See, for example, the possible attribution to Chrétien of the longer version of the tale found in *Guillaume*, discussed p. 142n4.

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. Adding *Guillaume*, *Robert*, and the putative earlier versions of *Richard* and *Jehan Paulus* to Jehan's *œuvre* acts as an encouragement for future scholars to bring these texts into conversation with each other and with those texts previously attributed to Jehan, opening up new avenues for literary analysis.

\* \* \*

As Jehan so often does in his own texts, we end 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 éclaire 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 generative AI reveal more than ever that the unique identity of an author is a fragile notion which can 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 bespoke, humanistic way, combining scientific techniques with literary understanding.

## BIBLIOGRAPHY

### Manuscripts

Avignon, Bibliothèque municipale, MS 615  
Besançon, Bibliothèque municipale, 588  
Carpentras, Bibliothèque Inguimbertaine, 106  
Grenoble, Bibliothèque municipale, MS 319  
London, British Library, MS Additional 15606  
Manchester, John Rylands University Library, MS French 6  
Paris, Bibliothèque nationale de France, MS Arsenal 2115  
Paris, Bibliothèque nationale de France, MS français 1555  
Paris, Bibliothèque nationale de France, MS français 1881  
Paris, Bibliothèque nationale de France, MS français 2198  
Paris, Bibliothèque nationale de France, MS français 5036  
Paris, Bibliothèque nationale de France, MS français 12483  
Paris, Bibliothèque nationale de France, MS français 12604  
Paris, Bibliothèque nationale de France, MS français 24432  
Paris, Bibliothèque nationale de France, MS nouvelle acquisition française 10721

### Printed Primary Sources

Allyn Williams, Charles, and Louis Allen. 1935. *The German Legends of the Hairy Anchorite, with Two Old French Texts of "La vie de saint Jehan Paulus."* University of Illinois Press.

Anonymous. ca. 1490–1500. *S'ensuyt le romant de Richart filz de Robert le diable qui fut duc de Normendie.* [Paris, Bibliothèque nationale de France, Rés. Y2-713.]

Bastin, Julia, and Edmond Faral, eds. 1959–1960. *Œuvres complètes de Rutebeuf.* 2 vols. Picard.

Blangez, Gérard, ed. 1979–1986. *Ci nous dit: Recueil d'exemples moraux.* 2 vols. Société des anciens textes français.

Breul, Karl, ed. 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*, 464–509. Niemeyer.

Buzzetti Gallarati, Silvia, ed. 1990. *Dit de Guillaume d'Engleterre.* Orso.

Chaurand, Jacques, ed. 1971. *Fou, dixième conte de la "vie des Pères," conte pieux du XIII<sup>e</sup> siècle.* Droz.

Conlon, Denis Joseph, ed. 1977. *Richard sans Peur: Edited from Le Roman de Richart and from Gilles Corrozet's Richart sans Paour.* University of North Carolina Press.

Denomy, Alex J., ed. 1956. "An Old French Version of the Julian Episode in the *Life of Saint Basil*." *Mediaeval Studies* 18: 105–24.

Dirickx-Van der Straeten, Hermine, ed. 1931. *La vie de saint Jehan Bouche d'or et la vie de sainte Dieudonnee, sa mère.* Vaillant-Carmanne.

Dows-Miller, Sebastian, ed. 2025a. "Le dit de menage: A Fourteenth-Century Set of Anti-Matrimonial, Anti-Clerical, Anti-English Dialogues." *Zeitschrift für romanische Philologie* 141 (2): 350–73.

- Eastman Manning, Helen, ed. 1929. *La vie de saint Thibaut: An Old French Poem of the Thirteenth Century*. Institute of French Studies.
- Fant, Carl, ed. 1887. *Légende de saint Grégoire: Rédaction du XIVe siècle*. Almqvist.
- Fudeman, Kirsten A., ed. 2008. "Restoring a Vernacular Jewish Voice: The Old French Elegy of Troyes." *Jewish Studies Quarterly* 15 (3): 190–221.
- Gaucher, Élisabeth, ed. 2006. *Robert le Diable: Édition bilingue*. Champion.
- Hamel, Anton Gerard van, ed. 1885. *Li romans de carité et Miserere du Renclus de Moiliens: Poèmes de la fin du XIIe siècle*. Vieweg.
- Jubinal, Achille, ed. 1839–1842. *Nouveau recueil de contes, dits, fabliaux et autres pièces inédites des XIIIe, XIVe et XVe siècles*. 2 vols. Pannier.
- Kastner, Léon-Émile, ed. 1906. "Les versions françaises inédites de la descente de saint Paul en enfer." *Revue des langues romanes* 49: 427–50.
- Koenig, V. Frederic, ed. 1955–1970. *Les miracles de Nostre Dame par Gautier de Coinci*. 4 vols. Droz.
- Långfors, Arthur, ed. 1932. *Notice des manuscrits 535 de la Bibliothèque municipale de Metz et 10047 des nouvelles acquisitions du fonds français de la Bibliothèque nationale suivie de cinq poèmes français sur la parabole des quatre filles Dieu*. Imprimerie nationale.
- Méon, Dominique Martin, ed. 1823. *Nouveau recueil de fabliaux et contes inédits, des poètes français des XIIe, XIIIe, XIVe et XVe siècles*. 2 vols. Chassieriau.
- Micha, Hugues, ed. 1971. "Une rédaction en vers de la vie de saint Sébastien." *Romania* 92: 405–19.
- 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 XIe et XIIe siècles*. 3 vols. Frère.
- Migne, Jacques-Paul, ed. 1845. *Sanctorum Zenonis et Optati, prioris Veronae, alterius Milevi episcoporum*. *Patrologiae cursus completus: Series latina*, 11. Migne.
- Munk Olsen, Birger, ed. 1978. *Dits en quatrains d'alexandrins monorimes de Jehan de Saint-Quentin*. Société des anciens textes français.
- Mussafia, Adolf, ed. 1870. *Über eine altfranzösische Handschrift der k. Universitätsbibliothek zu Pavia*. Hof- und Staatsdruckerei.
- Neufang, Alfred, ed. 1914. *Mitteilungen aus der Alexandrinerversion der chanson d'Ami et d'Amille*. Adler.
- Noomen, Willem, and Nico van den Boogaard, eds. 1983–1998. *Nouveau recueil complet des fabliaux*. 10 vols. van Gorcum.
- Paris, Gaston, and Léopold Pannier, eds. 1872. *La vie de saint Alexis: Poème du XIe siècle et renouvellements des XIIe, XIIIe et XIVe siècles*. Bibliothèque de l'École des hautes études.
- Raynaud, Gaston, ed. 1885. "Le clerc de Voudoi." *Romania* 51: 278–79.
- Short, Ian, ed. 2021. *Crestien's Guillaume d'Angleterre—William of England: An Edition and Annotated Translation*. University of Exeter Press.
- Tarbé, Prosper, ed. 1851. *Le tournoiement de l'Antéchrist par Huon de Méry*. Regnier.
- Toulmin Smith, Lucy, and Paul Meyer, eds. 1889. *Les contes moralisés de Nicole Bozon, frère mineur*. Didot.
- Türk, Monika, ed. 2000. *Lucidaire de grant sapientie: Untersuchung und Edition der altfranzösischen Übersetzung 1 des "Elucidarium" von Honorius Augustodunensis*. Niemeyer.
- Walberg, Emmanuel, ed. 1922. *La vie de saint Thomas le martyr par Guernes de Pont-Sainte-Maxence: Poème historique du XIIe siècle (1172–1174)*. Gleerup.

- Wallensköld, Axel, ed. 1907–1909. *Florence de Rome: Chanson d'aventure du premier quart du XIIIe siècle*. 2 vols. Didot.
- Weiske, Brigitte, ed. 1991–1992. *Gesta Romanorum*. 2 vols. Niemeyer.
- Westberg, Erik, ed. 1943. *La vie de saint Jean l'Évangéliste: Poème religieux du XIIIe siècle*. Appelberg.
- Wolter, Eugen, ed. 1879. *Der Judenknabe: 5 griechische, 14 lateinische und 8 französische Texte*. Niemeyer.
- Zink, Michel, ed. 1989–1990. *Œuvres complètes de Rutebeuf: Texte établi, traduit, annoté et présenté avec variantes*. 2 vols. Garnier.

## Secondary Works

- Argamon, Shlomo. 2018. "Computational Forensic Authorship Analysis: Promises and Pitfalls." *Language and Law* 5 (2): 7–37.
- Barbieri, Luca. 2024. "Introduction." In *L'auteur dans ses livres: Autorité et matérialité dans les littératures romanes du Moyen Âge*. Edited by Luca Barbieri et al., 7–16. Reichert.
- Barthes, Roland. 1968. "La mort de l'auteur." *Manteia* 5: 12–17.
- Beaussart, François-Jérôme. 2000. "Héroïsme ou sainteté: Les versions religieuses et profanes de l'histoire de Florence de Rome." *PRIS-MA* 16 (31): 3–30.
- Berlioz, Jacques. 1994. *Moines et religieux au Moyen Âge*. Seuil.
- , and Alain Guerreau. 2019. "La chair dans les recueils d'exempla méridionaux (XIIIe-XIVe siècles)." *Cahiers de Fanjeaux* 52: 211–33.
- Berthelot, Anne. 1991. *Figures et fonction de l'écrivain au XIIIe siècle*. Institut d'études médiévales.
- Brémond, Claude, Jacques Le Goff, and Jean-Claude Schmitt. 1982. *L'exemplum*. Brepolis.
- Brunel, Clovis. 1926. *La fille du comte de Pontieu: Nouvelle du XIIIe siècle*. Champion.
- Busby, Keith. 2011. "The Manuscripts of Marie de France." In *A Companion to Marie de France*. Edited by Logan E. Whalen, 303–17. Brill.
- Campbell, Ian. 2007. "Chi-Squared and Fisher-Irwin Tests of Two-by-Two Tables with Small Sample Recommendations." *Statistics in Medicine* 26 (19): 3661–75.
- Cerquiglini, Bernard. 1981. *La parole médiévale*. Minuit.
- . 1989. *Éloge de la variante: Histoire critique de la philologie*. Seuil.
- Cerquiglini-Toulet, Jacqueline. 2001. *Un engin si subtil: Guillaume de Machaut et l'écriture au XIVe siècle*. Champion.
- . 2006. "Portraits of Authors at the End of the Middle Ages: Tombs in Majesty and Carnavalesque Epitaphs." Translated by Darla Gervais. In *The Medieval Author in Medieval French Literature*. Edited by Virginie Greene, 157–71. Palgrave.
- . 2022. "L'invention du dit: Quand le dit devient un écrit." In *Le Dit du berceau au tombeau (XIIIe-XVe siècle)*. Edited by Isabelle Delage-Béland and Anne Salamon, 129–46. Garnier.
- 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 (XIIIe-XVe siècle)*. Edited by Isabelle Delage-Béland and Anne Salamon, 91–108. Garnier.
- Cohen, Jacob, et al. 2013. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. 3rd ed. Routledge.
- Corbellari, Alain. 2004. "Un problème de paternité: Le cas d'Henri d'Andeli. I. Arguments littéraires." *Revue de linguistique romane* 68: 47–56.

- Craig, Hugh. 2021. "Shakespeare and Authorship Attribution Methodologies." In *The Arden Research Handbook of Shakespeare and Textual Studies*. Edited by Lukas Erne, 225–43. Arden.
- . 2024. "Principal Components Analysis in Stylometry." *Digital Scholarship in the Humanities* 39: 97–108.
- Delage-Béland, Isabelle, and Anne Salamon, eds. 2022. *Le Dit du berceau au tombeau (XIIIe–XVe siècle)*. Garnier.
- Ducrot, Oswald. 1984. *Le dire et le dit*. Minuit.
- . 1989. *Logique, structure et énonciation*. Minuit.
- Eder, Maciej, Jan Rybicki, and Mike Kestemont. 2016. "Stylometry with R: A Package for Computational Text Analysis." *R Journal* 8 (1): 107–21.
- Faral, Edmond. 1910. *Les jongleurs en France au Moyen Âge*. Champion.
- Foucault, Michel. 1994. "Qu'est-ce qu'un auteur?" In *Dits et écrits*. Vol. 1. Edited by Daniel Defert and François Ewald, 789–821. 4 vols. Gallimard.
- Friedl, Jeffrey E. F. 2006. *Mastering Regular Expressions*. 3rd ed. O'Reilly.
- Gingras, Francis. 2016. "Introduction." In *Motifs merveilleux et poétique des genres au Moyen Âge*. Edited by Francis Gingras, 7–20. Garnier.
- . 2022. "Façons de dire: Le cas du *Ci nous dit*." In *Le Dit du berceau au tombeau (XIIIe–XVe siècle)*. Edited by Isabelle Delage-Béland and Anne Salamon, 37–52. Garnier.
- Gowans, Linda. 1990. "Guillaume d'Angleterre: Prologue and Authorship." *French Studies Bulletin* 10 (35): 1–5.
- Greene, Virginie. 2006a. "Introduction." In *The Medieval Author in Medieval French Literature*. Edited by Virginie Greene, 1–11. Palgrave.
- . 2006b. "What Happened to Medievalists after the Death of the Author?" In *The Medieval Author in Medieval French Literature*. Edited by Virginie Greene, 205–27. Palgrave.
- Grente, Georges, et al. 1994. *Dictionnaire des lettres françaises: Le Moyen Âge*. 2nd ed. Fayard.
- Gröber, Gustav. 1888–1902. *Grundriss der romanischen Philologie*. 2 vols. Trübner.
- Heidegger, Martin. 1975. "The Origin of the Work of Art." In *Poetry, Language, Thought*. Translated by Albert Hofstadter, 15–88. Harper.
- Hellgardt, Ernst. 1998. "Anonymität und Autornamen zwischen Mündlichkeit und Schriftlichkeit in der deutschen Literatur des elften und zwölften Jahrhunderts. Mit Vorbemerkungen zu einigen Autornamen der altenglischen Dichtung." In *Autor und Autorschaft im Mittelalter*. Edited by Elizabeth Andersen et al., 46–72. Niemeyer.
- Jauss, Hans Robert. 1970. "Littérature médiévale et théorie des genres." *Poétique* 2: 79–101.
- Kedar, Benjamin Z. 1973. "Toponymic Surnames as Evidence of Origin: Some Medieval Views." *Viator* 4: 123–30.
- Kienzle, Beverly Mayne. 2000. "Introduction." In *The Sermon*. Edited by Beverly Mayne Kienzle, 143–74. Brepols.
- Kleist, Wolfram. 1973. *Die erzählende französische Dit-literatur in "quatrains alexandrins monorimes"*. Buske.
- Lacy, Norris J. 1987. "Fabliaux and the Question of Genre." *Reading Medieval Studies* 13: 25–34.
- Lefèvre, Yves, Philippe Ménard, Robert Martin, and Marc Wilmet. 1976–1983. *Manuel du français du Moyen Âge*. 3 vols. SOBODI.

- Lote, Georges. 1949–1996. *Histoire du vers français*. 9 vols. Presses universitaires de Provence.
- Marchello-Nizia, Christiane. 1979. *Histoire de la langue française aux XIVe et XVe siècles*. Bordas.
- , et al., eds. 2020. *Grande grammaire historique du français*. 2 vols. De Gruyter.
- Marnette, Sophie. 1998. *Narrateur et point de vue dans la littérature française médiévale: Une approche linguistique*. Lang.
- . 2005. *Speech and Thought Presentation in French*. Benjamins.
- . 2006a. "La signalisation du discours rapporté en français médiéval." *Langue française* 149 (1): 31–47.
- . 2006b. "The Experiencing Self and the Narrating Self in Medieval French Chronicles." In *The Medieval Author in Medieval French Literature*. Edited by Virginie Greene, 117–36. Palgrave.
- Minnis, Alastair J. 1988. *Medieval Theory of Authorship: Scholastic Literary Attitudes in the Later Middle Ages*. Scholar.
- Mosteller, Frederick, and David L. Wallace. 1963. "Inference in an Authorship Problem." *Journal of the American Statistical Association* 58 (302): 275–309.
- Morawski, Joseph. 1939. "Mélanges de littérature pieuse. II. Les miracles en quatrains alexandrins monorimes: Jean de Saint-Quentin." *Romania* 65: 327–58.
- . 1940. "Mélanges de littérature pieuse. II. Les miracles en quatrains alexandrins monorimes (suite et fin)." *Romania* 66: 505–29.
- Munk Olsen, Birger. 1963a. "Étude et édition des *dits* en quatrains alexandrins monorimes du manuscrit français 24432 de la Bibliothèque nationale." PhD diss., École pratique des hautes études, Paris.
- . 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*. Edited by Karl D. Uitti, 144–65. Munksgaard.
- Naetebus, Gotthold. 1891. *Die nicht-lyrischen Strophenformen des Altfranzösischen*. Hirzel.
- Nichols, Stephen G. 2006. "The Medieval 'Author': An Idea Whose Time Hadn't Come?" In *The Medieval Author in Medieval French Literature*. Edited by Virginie Greene, 77–101. Palgrave.
- Nykrog, Per. 1957. *Les fabliaux: Étude d'histoire littéraire et de stylistique médiévale*. Munksgaard.
- Pearson, Karl. 1895. "Notes on Regression and Inheritance in the Case of Two Parents." *Proceedings of the Royal Society of London* 58: 240–42.
- Polo de Beaulieu, Marie-Anne. 2019. *Éducation, prédication et cultures au Moyen Âge: Essai sur Jean Gobi le Jeune*. Presses universitaires de Lyon.
- Reenen, Pieter van, 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 VIIIe Colloque international sur le moyen français*. Edited by Bernard Combettes and Simone Monsonégo, 595–612. Didier.
- Rehr, Jean-Paul, and Marie-Anne Polo de Beaulieu. 2021. "Thesaurus exemplorum Medii Aevi: une base de données collaborative sur les exempla médiévaux." *Humanités numériques* 4, <https://doi.org/10.4000/revuehn.2630>.
- 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*. Edited by Sarah Baudelle-Michels et al., 729–39. Champion.

- Rubin, Miri. 1999. *Gentile Tales: The Narrative Assault on Late Medieval Jews*. Yale University Press.
- Ruxton, Graeme D. 2006. "The Unequal Variance t-test is an Underused Alternative to Student's t-test and the Mann-Whitney U Test." *Behavioral Ecology* 17 (4): 688–90.
- 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*. Edited by Adrian P. Tudor and Kristin L. Burr, 107–20. University Press of Florida.
- 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), <https://doi.org/10.21825/aj.v2i1.760>.
- Smirnova, Victoria. 2022. "Raconter des histoires dans une communauté soumise au silence: les Cisterciens et leurs *exempla* (XIIe-XIIIe siècles)." *L'atelier du Centre de recherches historiques* 25, <https://doi.org/10.4000/acr.26213>.
- Stout, Julien. 2025. *L'auteur retrouvé: L'avènement des premiers recueils à collections auctoriales de langue française au Moyen Âge central*. Droz.
- Stowell, William A. 1908. *Old French Titles of Respect in Direct Address*. Furst.
- Taylor, Larissa. 2000. "French Sermons, 1215–1535." In *The Sermon*. Edited by Beverly Mayne Kienzle, 711–58. Brepols.
- Tubach, Frederic C. 1969. *Index Exemplorum: A Handbook of Medieval Religious Tales*. Suomalainen Tiedekatemia.
- Wallensköld, Axel. 1907. *Le conte de la femme chaste convoitée par son beau-frère: Étude de littérature comparée*. Société de littérature finnoise.
- Welch, Bernard Lewis. 1947. "The Generalization of 'Student's' Problem when Several Different Population Variances are Involved." *Biometrika* 34 (1–2): 28–35.
- Welter, Jean Thiébaut. 1927. *L'exemplum dans la littérature religieuse et didactique du Moyen Âge*. Occitania.
- Wenzel, Siegfried. 2016. *Medieval "Artes praedicandi": A Synthesis of Scholastic Sermon Structure*. University of Toronto Press.
- Zimmerman, Michel. 2001. "Ouverture du colloque." In *Auctor et auctoritas: Invention et conformisme dans l'écriture médiévale*. Edited by Michel Zimmerman, 7–14. École des Chartes.
- Zufferey, François. 2004. "Un problème de paternité: Le cas d'Henri d'Andeli. II. Arguments linguistiques." *Revue de linguistique romane* 68: 57–78.
- Zumthor, Paul. 1972. *Essai de poésie médiévale*. Seuil.

## Digital Resources and Software

- Anthony, Laurence. 2023. *AntConc*. Waseda University. <https://www.laurenceanthony.net/software/antconc/>, accessed October 26, 2022.
- ATILF. 2020. *Dictionnaire du moyen français*. <http://www.atilf.fr/dmf>, accessed March 9, 2026.
- Brun, Laurent. 2023. *ARLIMA: Archives de littérature du Moyen Âge*. <https://www.arlima.net/>, accessed March 9, 2026.
- Dows-Miller, Sebastian, ed. 2025b. *The Texts of BnF fr. 24432: A Cumulative, Work-In-Progress Digital Edition*. <https://dowsmiller.github.io/fr24432>, accessed January 17, 2026.
- Guillot-Barbance et al. 2022. *Base de français médiéval*. ENS de Lyon. <http://bfm.ens-lyon.fr/>, accessed March 9, 2026.

- Lavrentiev, Alexei. 2023. *Logiciel RNNTagger*.  
<http://bfm.ens-lyon.fr/spip.php?article355>, accessed October 5, 2023.
- Lexical Computing. 2023. *Sketch Engine*. Brno: Lexical Computing.  
<https://www.sketchengine.eu/>, accessed March 9, 2026.
- OpenAI. 2022a. *ChatGPT*. <https://chat.openai.com/chat>, accessed March 9, 2026.
- . 2022b. *Introducing ChatGPT*. <https://openai.com/blog/chatgpt>, accessed August 12, 2023.
- Pierreville, Corinne (ed.). 2024. *Corpus FABLIAUX-V0-3*. <http://txm.bfm-corpus.org>, accessed March 9, 2026.
- Polo de Beaulieu et al. 2026. *À propos de ThEMA*. <https://thema.humA-num.fr/about/>, accessed March 9, 2026.
- R. Core Team. 2025. *R: A Language and Environment for Statistical Computing*. R. Foundation for Statistical Computing. <https://www.R-project.org/>, accessed March 9, 2026.
- Sentjens, Cedric. 2023. *Rutebeuf.be*. <http://www.rutebeuf.be>, accessed March 9, 2026.
- 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 September 7, 2023.
- . 2019. *RNNTagger*. Ludwig-Maximilians-Universität. <https://www.cis.uni-muenchen.de/~schmid/tools/RNNTagger/>, accessed October 6, 2023.
- TEI Consortium, eds. 2025. *Guidelines for Electronic Text Encoding and Interchange*. <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html>, accessed March 9, 2026.
- World Wide Web Consortium (W3C). 1999. *XSL Transformations (XSLT) Version 1.0*. <https://www.w3.org/TR/1999/REC-xslt-19991116/>, accessed March 9, 2026.



## INDEX OPERUM

### **J-corpus**

- Abesse*, 6–7, 9, 22–24, 46, 55, 58, 83,  
89–90, 92–93, 98, 115–17, 120–23,  
131, 133, 141
- Anelés*, 8–9, 11, 31, 33–35, 38, 50, 55,  
57, 63, 65–66, 68–69, 83, 90, 98,  
121, 123, 127, 131–35, 136, 142–44
- Beguine*, 7, 9, 35, 41–42, 71, 89–90, 98,  
121, 123, 129, 130–31, 133, 142
- Buef*, 7, 9, 11–12, 31, 33–35, 38, 44, 46,  
50–51, 55, 65–66, 71, 78, 90, 92,  
97–98, 121–23, 127, 131, 133–34,  
14–43, 147, 155
- Chanoines*, 7, 9, 31, 46, 50–51, 69,  
71–75, 98, 116–17, 121, 123, 129,  
131, 133, 141, 147–48
- Chevalier et escuier*, 5–7, 9, 11–12,  
29–30, 34, 39, 42, 55, 61, 63, 69,  
88–90, 98, 116–18, 121–23, 129,  
131, 133, 141
- Chevalier hermite*, 7, 9, 65–66, 78–79,  
90, 98, 121, 123, 129, 131, 133,  
142, 146
- Cordouanier*, 7, 9–10, 30, 55, 57, 63, 65,  
72–73, 90, 98, 121–23, 130–31,  
133–34, 142
- Deux chevaliers*, 7, 9, 50, 65, 78, 88, 90,  
98, 116, 121, 123, 131, 133, 142,  
148–49
- Eaue et vergier*, 7, 9, 50, 51, 71–72, 90,  
98, 121, 123, 131, 133, 141–42,  
147, 150
- Enfant rosti*, 7, 9, 39, 50–51, 65, 89–90,  
98, 121–23, 131, 133, 142, 145
- Enfant sauva mere*, 7, 9, 50, 55, 71–72,  
90, 98, 121, 123, 131, 133, 142,  
145, 149
- Florence*, 6–9, 14, 16, 30–31, 33–34, 38,  
44, 65, 68–69, 72, 78, 83, 85, 89–90,  
92, 97–99, 120–23, 127, 130–35,  
136, 142, 147–48, 152–55
- Juitel*, 7, 9, 34, 68, 72–73, 90, 98,  
121, 123, 129, 131, 133, 141
- Merlin*, 6–7, 9, 14, 50, 68–69, 71, 89–90,  
98, 121, 123, 130–33, 141
- Mescreant*, 7, 9, 46, 50, 61, 63, 65, 68,  
72–73, 78, 89–90, 97–98, 121, 123,  
131–33, 142, 147, 149–52
- Narbonne*, 7, 9, 11–12, 43–44, 46,  
50–51, 55, 65, 68, 70–72, 78,  
90–91, 98, 116, 121, 123, 131, 133,  
142, 145
- Pain*, 7, 9, 49, 89–91, 98, 121, 123, 129,  
131, 133, 142, 147–49
- Pecherresse*, 7, 9, 30, 39, 43, 46, 65,  
73–75, 78, 90–91, 98, 121, 123,  
131, 133, 141
- Pommes*, 7, 9, 37, 50, 55, 57, 61, 63,  
68, 78–79, 90, 98, 116, 121–23,  
131, 133, 142
- Povre chevalier*, 7, 9, 37, 71, 89–90, 98,  
115–18, 121, 123, 131–33, 142
- Respon*, 7, 9, 41–42, 50, 55, 69, 89–90,  
98, 116, 121, 123, 129, 131, 133,  
141–42
- Romme*, 7, 9, 39, 44, 50–51, 63, 90–92,  
96–98, 116, 121–23, 131, 133, 141,  
147
- Sauveur*, 6–7, 9, 16, 22, 45, 55, 58, 78,  
83, 90, 98, 121–23, 129, 131, 133,  
142, 147, 150

**A-corpus**

*Alexis*, 8–9, 11, 44, 63, 93–94, 100, 103, 119, 136  
*Barbe*, 35, 72, 80, 93–94, 136  
*Basile*, 8, 33, 35, 44–45, 70, 93–94, 100, 102, 136, 153  
*Confesseur*, 33, 35, 72, 88, 93–94, 100, 135–36, 151  
*Cristofle*, 44, 56, 94, 100, 136  
*Déz*, 35, 72, 94, 100, 135–36  
*Dieu d'Amours*, 26, 56, 93–94, 99–100, 135  
*Dieudonnée*, 8, 35, 44, 56, 66, 93–94, 100, 136  
*Grégoire*, 8, 35, 56, 70, 93–94, 100, 136, 142  
*Guillaume*, 4, 9, 11, 13–14, 27, 29, 34–35, 37–38, 53, 56, 58, 60, 67–68, 72–75, 77, 83, 85–87, 93–95, 99–101, 107, 110–11, 113, 126–28, 135–39, 141–42, 145–49, 151, 155–57, 160–61  
*Jehan Ewangeliste*, 11, 35, 44–45, 56, 66, 94, 100, 136  
*Jehan Paulus*, 4, 8, 13, 16, 33–35, 37–38, 44, 61, 63–64, 67, 72–75, 77, 84–87, 93–95, 99–101, 107, 110–11, 113, 116–17, 126–27, 135–39, 142, 146, 156–57, 160  
*Leu*, 33, 44, 51, 94, 100, 136, 142  
*Mauvais riche home*, 26, 94, 100, 136  
*Quatre sereurs*, 11, 33, 35, 51, 93–94, 100, 136  
*Richard*, 4, 8, 11, 13–16, 29, 34–35, 37–38, 44, 60–61, 63–64, 67, 72–75, 77, 84–87, 93–95, 99–101, 107, 110–11, 113, 126–28, 135–39, 142, 146, 155–57, 160  
*Robert*, 4, 8, 9, 13–15, 27, 29, 34–35, 37–38, 56, 58, 60–61, 63–64, 66, 68, 74–75, 77, 83, 85–87, 93–95, 99–101, 107, 110–13, 126–28, 135–39, 141–42, 146–48, 155–57, 160

*Roy amie*, 33, 35, 94, 100, 136  
*Sebastien*, 33, 35, 63, 93–94, 100, 136, 144  
*Servais*, 8, 35, 94, 100, 136  
*Thibaut*, 35, 43, 45, 88, 94, 100, 136, 145–46  
*Thomas*, 35, 93–94, 100, 135, 144  
*Troyes*, 35, 92, 94, 99–100, 136  
*Vieillards*, 35, 88, 93–94, 99–100, 136  
*Vision*, 11, 27, 35, 88, 94, 100, 136  
*Yves*, 26, 93–94, 100, 135

**B-corpus**

*Fauvel, Le roman de*, 80  
*Passion, La*, 80  
*Passion, Le livre de la*, 80  
*Prunier, Le dit du*, 80  
*Ressource de la chrestienté*, 72  
*Tyolet, Le lai de*, 151

**P-corpus**

*Enfant qui mist l'anel que s'amie li avoit donné ou doit a l'ymage de Nostre Dame, Le dit de l'*, 41  
*Lucidaire*, 82  
*Tournoiement d'Antecrist et d'euls de son barnage, Le dit du*, 33

**R-corpus**

*Brichemer, De*, 102  
*Cordeliers, Le diz des*, 102  
*Elysabel, La vie sainte*, 34  
*Jacobins, Des*, 102  
*Puille, La chanson de*, 102