Chapter 1

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,[1] 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

1 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.

DOI: 10.17302/ogkl54866 © Arc Humanities Press

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.[2]

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,[3] 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 ± 0.5 and 1, and weak when it is between ± 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

```r
#randomised p_value tests for target vs reference corpus
p_value_rand <- function(
    observed_tar = integer(),
    total_tar = integer(),
    observed_ref = integer(),
    total_ref = integer(),
    sample_size = 1000000) {

    sum_observed <- observed_tar + observed_ref #sum of observed instances
    sum_total <- total_tar + total_ref #sum of total instances

    #null rate: expected rate if there is no difference between a and b
    null_rate <- sum_observed / sum_total

    #observed difference between probabilities
    observed_prob_dif <- (observed_ref / total_ref) - (observed_tar / total_tar)

    #generate random samples, counting observed instances
    ref_rand <- rbinom(sample_size, total_ref, null_rate)
    tar_rand <- rbinom(sample_size, total_tar, null_rate)

    #work out differences in probability between two samples
    rand_prob_difs <- (ref_rand / total_ref) - (tar_rand / total_tar)

    #make differences in probability absolute
    abs_rand_prob_difs <- abs(rand_prob_difs)
    abs_observed_prob_dif <- abs(observed_prob_dif)

    #calculate the p value (the proportion of the random sample for which
        #the difference in probability exceeds the observed probability)
    p_value <- mean(abs_rand_prob_difs >= abs_observed_prob_dif)

    p_value
}
```

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.[4] 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: *öir* and *ouïr*, despite being forms of the same verb, likely due to discrepancies in the training data.

An alternative approach would be to run the data through a more generalist AI natural language model, such as ChatGPT (OpenAI 2022a). Table 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.[5] 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.[6]

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.[7] 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

7 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.[8]

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.[9] 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 home* (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 Inguimbertine, 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.[10] 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.[11]

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).[12] 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.