Digital Editing and Linguistic Analysis
The First Redaction of the
Histoire ancienne jusqu’à César

Simone Ventura

Abstract
This essay examines how the availability of a set of digital tools (including a complete digital edition and an instrument for textual comparison through a complex textual tradition) can help us in the study of ITS language. This paper is based on the Histoire ancienne jusqu’à César, an early 13th French universal chronicle. This is the core text of The Values of French Language and Literature in the European Middle Ages, an ERC-Advanced Grant based at King’s College London (PI, Simon Gaunt). The digital edition of the complete text of the Histoire ancienne is paralleled by the availability of a digital tool, Alignment, that maps the contents of the Histoire ancienne in its rich manuscript tradition. At the same time, we would not have realized the “competition-behavior” in language among the manuscripts without Alignment and the systematic study of the textual tradition. Section §2 illustrates how the digital edition and Alignment have been used as grounds for the study of the language of the Histoire ancienne’s textual tradition.

The Values of French Language and Literature in the European Middle Ages is an ERC-Advanced Grant project focused on The Histoire ancienne jusqu’à César (from here on Histoire ancienne), an early thirteenth-century French prose work that provides us with the core textual tradition and is the case study and testing ground for the project.1

1. While we focus on thirteenth-century prose, as we shall see, we understand Old French as stretching from the ninth to the fourteenth century. It goes without saying that fifteenth-century manuscripts passing down thirteenth-century texts (recentiores) are included. The Histoire ancienne is considered to be one of the earliest and most important works in Old French prose. However, thirty pieces in octosyllabic couplets are included in this work. For the most part
The *Histoire ancienne* was compiled in North-Eastern France during the first quarter of the thirteenth century. It then circulated widely throughout Europe and the Mediterranean. The textual tradition of this work encompasses the Outremer territories, Italy, and France. Throughout its first two centuries of life, the *Histoire ancienne* appeared in two major textual forms. The so-called first redaction is in fact the original version of the work.\(^2\) The second redaction is a structural and stylistic modification of the first redaction.\(^3\) In this paper I will show how the complete digital edition and the development of a specific tool for manuscript comparison, the Alignement, informed my approach to the language of the *Histoire ancienne*'s textual tradition.\(^4\) To exemplify this approach, I will consider the syntax of hypothetical comparative clauses (type: “as if it” followed by an inflected verb). We will see how the availability of a digital edition and the Alignement allowed me to grasp linguistic dynamic trends throughout the tradition. On this bases, different inferences could be drawn both at linguistic and textual levels.

In its current form, the first redaction of the *Histoire ancienne* is a universal history, dealing with human events from the creation to Caesar's...
military campaign in Gaul. The first redaction survives in over eighty manuscripts that have been copied and ‘edited’ in the Holy Land (Acre), Italy and France. It is in this version that the *Histoire ancienne* became one of the most successful companions to ancient history in the vernacular.5

By presenting the text and transmission of the *Histoire ancienne* as both a case study and a testing ground, we intend to make three points. The first is that the *Histoire ancienne* has strong potential to become a methodological paradigm. It raises various broader, general issues about the history and purposes of thirteenth-century French prose, about the meaning and rhetoric of historical writing in late medieval Europe (and beyond), on the “values” of French as a historical language and a language for historiography.6 Secondly, the *Histoire ancienne* poses its own specific editing problems that we have dealt with at three different levels: through digital manuscript editing, through the investigation of its textual transmission, and through the analysis of its language.7 The third is that the language of the *Histoire ancienne* cannot be disassociated from the ‘individuals’ who used it. This meant for me to approach the study of the language of the *Histoire ancienne* in the manuscripts that we possess (as opposed to the edited text alone). My perspective is twofold: as individuals, manuscripts show specific linguistic features that deserve attention per se; as elements of a textual and linguistic continuum (tradition), the similarities that the manuscripts show are taken along the differences that break the linguistic continuum down into discrete units.8

5. The second redaction of the HA was compiled at the beginning/during the second quarter of the fourteenth century at the Neapolitan court of the Angevins. The second redaction of the *Histoire ancienne* is characterized by the omission of the Biblical and Alexander sections, while it preserves the ‘matters’ of Thebes, Aeneas, and Rome, which are completed by the insertion of the fifth mise en prose of the Roman de Troie.

6. Within TVOF, this aspect is developed by the research projects of Maria Teresa Rachetta and Heny Ravenhall. Hannah Morcos is responsible for the editorial seam of the project.

7. As Michele Barbi (1938, x–xi) put it: “ogni testo ha il suo problema critico, ogni problema la sua soluzione, e che quindi le edizioni non si fanno su modello e, per così dire, a macchina”. It is maybe not without irony that — responding precisely to this principle — current digital editions, including ours, are based on digital modeling and workflow development through machine processing.

8. See below §2. For the dialectical continuum / parts in linguistic analysis see Weinreich 1954. As for manuscripts as individuals bearing crucial linguistic information, see Fleischman 2000, 34–35. In a private gloss to a quotation
Each of these three points is taken into account in the division of labor internal to the TVOF team (see below §1). The relevant aspects of the *Histoire ancienne* are approached from three different angles: digital editing of two manuscripts selected for their inherent relevance and role at the heart of the *Histoire ancienne*’s tradition, the development and implementation of **Alignment**, a digital tool for the thorough study of textual transmission, and related forms of language analysis along the lines that I am going to clarify below.

The contents of the paper are structured as follows. In § 1, I will briefly describe the three main digital outputs of the project as well as the approach and the digital workflow underpinning our editorial project and our analysis of the textual tradition. In § 2, I will draw on a series of examples taken from the edition of the *Histoire ancienne* tradition to illustrate my approach to language and show how digital editing of one manuscript along with the study of the textual tradition via **Alignment** constitute the backbone of the linguistic study. In the conclusion (§3), I will return to the considerations and data presented in §2 to raise some questions that remain unanswered in this paper and that will be at the center of my future work on thirteenth and fourteenth-century French.

### 1. The *Histoire ancienne jusqu’à César*: digital scholarly outputs

*The Values of French* project is producing the following three main digital outputs.

1) Led by Hannah Morcos (KCL), our *Digital edition* of the *Histoire ancienne jusqu’à César* provides the complete semi-diplomatic and interpretive editions of two of the most important manuscripts passing down the first and second redactions of this work respectively: Paris, BNF, f. fr. 20125 (= from Hjemslev on the objective and independent existence from the individuals of the linguistic system, another medievalist sensitive to language, Franca Brambilla Ageno, argued: “Si potrebbe obiettare che il linguaggio non è un oggetto ma un ‘fare’ (secondo certe regole) e, come qualunque ‘fare’, non ha luogo indipendentemente dagli individui che ‘fanno’. Si potrà ammettere l’esistenza oggettiva (in realtà la conoscenza comune) delle regole di questo ‘fare’”. Ageno’s annotation is on the margins of her own copy of the Italian edition of 1964 Lászlo Antal’s book *Content, Meaning, and Understanding* (see Canova 2015, 103).
Our editorial work is to provide the first complete text of the first and second redactions of the *Histoire ancienne* this work. At present, only partial editions of this work are available, the text remaining largely unedited and underexplored.\(^9\)

The *Histoire ancienne* is a very long text: 410,000 words in fr20125 and ca. 300,000 in MS Royal 20 D 1. Secondly, the textual tradition is complex. Particularly in its first redaction, the *Histoire ancienne* is passed down by a very rich and probably contaminated tradition. The Parisian manuscript passes down the longest and most complete version of the work, including the over thirty verse “moralizations” interspersed in the different sections of the *Histoire ancienne*. This manuscript constitutes the *manuscrit de base* of all extant partial editions. Its availability will allow the scholarly community to have access to the whole text of the first redaction. Moreover, the *fr20125* is a unique linguistic monument deserving a comprehensive study in its own right.

With regard to the second redaction, we have a completely different case. As recently proven by Luca Barbieri, the MS Royal 20 D I, the second codex that we are editing, not only represents the oldest surviving copy of the *Histoire ancienne*, it most probably is the copy on which all the other extant manuscript copies depend. In the stemmatic jargon, the MS Royal 20 D I is a surviving archetype.

2) The second digital output is the *Alignment* tool (http://www.tvof.ac.uk/histoire-ancienne/alignment, last accessed 6 January 2019). This tool is producing the first complete mapping of the *Histoire ancienne*’s textual contents throughout the tradition. The analysis of the textual tradition is crucial to understand the dynamic and the history of this text. For this purpose, *Alignment* maps the contents of the manuscripts of the tradition and allows for their comparison at a structural and macro-textual level. It is thanks to this tool that we have made substantial progress in our knowledge of the textual tradition of the *Histoire ancienne*. This work is the basis for our rationalization of the stemmatic relations between the very high number of manuscripts of the first and second redactions of the *Histoire ancienne*.

\(^9\) Hannah Morcos’s semi-diplomatic edition of Paris, BnF, f. fr. 20125 is available online: http://www.tvof.ac.uk/textviewer/ (last accessed 6 January 2019). The edition is in progress. The section *Eneas* is available also in an interpretative version with commentary.

\(^{10}\) The most comprehensive study of the *HA*’s tradition is still MEYER 1885. For a critical synthesis of the current state of affairs in relation to the *HA*, see TRACHSLEIN 2013. For an interpretation of the status of history (and fiction) in the *HA* and in its manuscript tradition, see GAUNT 2016.
ancienne. It will provide the basis for further progress on the editorial work on the text of the *Histoire ancienne*. At the moment, *Alignment*, already available online, records and represents visually some selected textual features of five manuscripts.\textsuperscript{11} The contents are give and id number, according to the paragraph division of Paris, BnF, f. fr. 20125. The information displayed in *Alignment* includes rubrics, prose and verse form, material and non-material lacunae. In the coming months, we will add the description of new features, including paratextual (e.g., large initials and presence/disposition of miniatures) and new manuscripts. However, this first version already demonstrates the rationale of *Alignment* and offers three different visualizations of the data, each designed for a different purpose:

- **Table**: to locate a paragraph in one of the manuscripts or compare the contents or rubrics of a short range of paragraphs. The sequence of paragraphs will appear in a classic tabular presentation.
- **Bars**: to have a synthetic view of larger textual ranges, such as entire sections or the whole text. It is also the most suitable solution for individuating patterns of variation.
- **Column**: to both access to details concerning single paragraphs and an overall view.

Users can choose which manuscript(s) and which narrative units of the *Histoire ancienne jusqu’à César* they want to be displayed. Finally, the information can be displayed according to the following parameters:

- **locus** = localization of each paragraph in manuscripts (“folio” and column);
- **rubric** = text of the rubrics; displacement of rubrics and additional ones, if any; where there is no information about rubrics, it means that the paragraph is copied continuously after the preceding one. The indication “Rubric: null” appears in two cases:

\textsuperscript{11} Paris, BnF, f. fr. 20125; London, BL, Royal 20 D I (the alignment of Royal includes *Prose 5*, a section which is exclusive to the second redaction); London, BL, Additional 15268; London, BL, Additional 19669; Paris, BnF, f. fr. 17177. In the coming months also the data concerning Paris, BnF, f. fr. 686 will be available. See below §2 (and Rachetta 2018 and Rachetta forthcoming) for further data concerning the relevance of these manuscripts within the textual tradition of the *Histoire ancienne*. 
◊ when there is a blank space for a rubric in the manuscript, but no rubric,
◊ when the beginning of a paragraph is part of a lacuna, hence the rubric may have been present originally;
• verse = verse form of a paragraph; it is also indicated if the verses are lineated (abbreviated to “lin”) or continuously copied (abbreviated to “cont”);
• variation = occurrence of a material lacuna (total or partial);
• note = additional information necessary to understand the configuration of the alignment (i.e., information about non-material lacunae, displacement of folios after the manuscript was copied).

Alignment is integrated with the digital edition of the *Histoire ancienne jusqu’à César*: users can directly access the text of each paragraph in Paris, BnF, f. fr. 20125 and London, BL, Royal 20 D I from the three visualizations by clicking on “Read” or the relevant bar.

3) The third major digital output concerns the complete lemmatization of the text of the manuscripts that we are editing. We are doing this in collaboration with Stephen Dörr and Markus Husar of the *Dictionnaire Étymologique de l’Ancien Français* (DEAF). We are developing a research tool that will allow searches by lemmas or inflected forms. The lemmas are standardized according to Tobler–Lommatsch (1925–1976) and, whenever necessary, modified according to the guidelines of the *Dictionnaire Étymologique de l’Ancien Français*.

From a technical standpoint, the edition and lemmatization are based on a unique digital workflow. The text of each of the manuscripts that we are editing is saved into multiple XML files. The TEI schema, designed by Paul Caton (King’s Digital Lab [KDL], King’s College London), semantically captures the way the text physically manifests itself on the charta and how it is interpreted by the editor. Working on multiple files helped distribute the editorial work among team members (without risk of clashes) and keeps the files small and manageable. The changes are made offline using an XML editor that validates the format of the files and immediately copies the content to a shared web space. Every two hours a program written by Geoffroy Nöel (KDL, King’s College London) and running on the web server copies the files from Dropbox, links them into a single large file and expands all the editorial short hands. The project team can then use the Text Viewer on the website to preview their latest changes (http://www.tvof.ac.uk/textviewer/, last accessed 12 September 2018). The Text Viewer is able to retrieve any portion of the text from the aggregated file, convert
it to HTML on the fly, and render it in the browser. This fragmentation of the text into small bits not only makes the navigation more comfortable and responsive but also allows easy sharing of links to any specific location in the text. The key benefit of this workflow is that the preview the researchers use to check their latest changes is totally consistent with the rendering which will eventually be offered to the end-users on the public website. Encoding issues can therefore be corrected early and directly in the source files and workflow bugs reliably reported to a KDL developer for further analysis and resolved without blocking or disrupting the editorial process.

The second part of the automated conversion workflow is related to the lemmatization of the text. The texts of the two edited manuscripts are fully lemmatized, an innovative and significant contribution to the field of digital editions of medieval French texts. The aggregated files are fully “tokenized” and a “keyword in context” (KWIC) is produced from it. The KWIC file is fed into Lemming, the online lemmatization tool elaborated by Marcus Husar and Stephen Dörr for the *Dictionnaire Étymologique de l’Ancien Français* at Heidelberg. All the lemmatization information will eventually be exported from Lemming to be incorporated into the text: the Text Viewer and a new search page on the site will let users search the text by lemma or form. Care was taken by the team for this part of the workflow to accept minor changes in the input texts to be reprocessed by Lemming without losing any data. The availability of the complete text of fr 20125 and the advanced elaboration of both Alignment and the lemmatization tools were crucial for the linguistic approach that I adopted and that I will illustrate in the next section.

2. Textual and linguistic variation

In this section, I will present my approach to the study of the language of the *Histoire ancienne*. The main research question is: How is textual variation linked to linguistic change? What is at stake is how medieval scribes negotiated their position between their own linguistic competence, the perception that they had of the features of the language reflected in their exemplar(s), and the communicative goals set up by the text that they were transcribing (*De Roberto* 2014, 494–5). The objectives of this approach are twofold: first, to acquire a better understanding of the language of the manuscripts that we are editing; second, to grasp the link between textual variation and linguistic constraints. Textual variants respond to a number
of linguistic and extra-linguistic factors. From a linguistic standpoint, textual variation takes place under certain conditions and within delimited margins that correspond to the “rules” of the linguistic system or subsystem. My approach will focus on these linguistic conditions and margins. The hypothesis to be tested is that this may have implications for our editorial practice and could provide us with new evidence about some of the major changes that the French language underwent over the thirteenth and fourteenth centuries.

I will focus on the textual transmission of the first redaction. I will compare some relevant linguistic features of the manuscript of the first redaction that we are editing, Paris, BnF, f. fr. 20125 (= fr20125), with those of a sample of manuscripts relevant from the point of view of their mutual textual relationships and of their geographic and historical context of production.\(^\text{12}\) Given the length of the text of the *Histoire ancienne*, the number of manuscripts of the first redaction, and the range of potentially interesting linguistic features, I have worked on a restricted corpus of manuscripts and on a selection of relevant linguistic features. With regard to the manuscripts of the *Histoire ancienne*, I relied on previous work on the textual tradition updated with Maria Teresa Rachetta’s research undertaken in the framework of the TVOF project.\(^\text{13}\) In so doing, I picked a set of manuscripts representative of all the acknowledged manuscript families and branches of the textual tradition of the *Histoire ancienne* (see Table 1).

As the following observations on the language of the *Histoire ancienne* are based on these manuscripts and their grouping, I will provide a brief synopsis of the main features of the acknowledged manuscript families and their mutual relationships:\(^\text{14}\)

\(^{12}\) For the manuscript references see Table 1 below. For a similar approach, see Buridant 2000 and Schøsler and Völker 2014. Zinelli 2011 and 2016b applies current approaches in contact linguistics to the understanding of the linguistic dimension of *scripta* and to textual reconstruction.


\(^{14}\) What follows relies on the substantial contribution to the rationalization and knowledge of the textual tradition of the *Histoire ancienne* provided by Zinelli 2016a (especially with regards to fr686) and Rachetta forthcoming (especially in relation to the Abridged version and the structure of Vienna).
<table>
<thead>
<tr>
<th><strong>MS</strong></th>
<th><strong>Date</strong></th>
<th><strong>Place</strong></th>
<th><strong>Siglum</strong></th>
<th><strong>Group</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Paris, BnF, f. fr. 686</td>
<td>13th c. ex.</td>
<td>Italy</td>
<td>fr686</td>
<td></td>
</tr>
<tr>
<td>Paris, BnF, f. fr. 9682</td>
<td>14th c. in.</td>
<td>France</td>
<td>fr9682</td>
<td>Acre Group</td>
</tr>
<tr>
<td>Brussels, KBR, MS 10175</td>
<td>13th c. ex.</td>
<td>Acre</td>
<td>B10175</td>
<td></td>
</tr>
<tr>
<td>Dijon, BM, MS 562</td>
<td>13th c. ex.</td>
<td>Acre</td>
<td>D562</td>
<td></td>
</tr>
<tr>
<td>London, British Library, Additional MS 15268</td>
<td>13th c. ex.</td>
<td>Acre</td>
<td>Add15268</td>
<td></td>
</tr>
<tr>
<td>Paris, BnF, f. fr. 20125</td>
<td>13th c. second ½</td>
<td>North-Eastern France/French Flanders?</td>
<td>fr20125</td>
<td></td>
</tr>
<tr>
<td>Rennes, Bibliothèque Municipale, MS 2331</td>
<td>15th c.</td>
<td>Brittany</td>
<td>Rennes</td>
<td></td>
</tr>
<tr>
<td>Vienna, ÖNB, MS 2576</td>
<td>14th c.</td>
<td>Venice</td>
<td>Vienna</td>
<td></td>
</tr>
<tr>
<td>London, British Library, Additional MS 19669</td>
<td>13th c. second ½</td>
<td>North-Eastern France/French Flanders</td>
<td>Add19669</td>
<td>Abridged Version</td>
</tr>
<tr>
<td>The Hague, MS 78D47</td>
<td>13th c. second ½</td>
<td>North-Eastern France/French Flanders</td>
<td>Hague78</td>
<td></td>
</tr>
<tr>
<td>Paris, Bibliothèque Nationale de France, f. fr. 17177</td>
<td>1280–1300</td>
<td>Soissons (see Giannini 2016)</td>
<td>fr17177</td>
<td></td>
</tr>
</tbody>
</table>

1. A different convention in De Visser–van Terwisga 12–14. Rachetta forthcoming clarifies the structure and the nature of the editorial project of the Vienna manuscripts. She designates by Vienna1 the sections of the manuscript closest to fr20125.

2. Codex fr9682 is an early fourteenth-century manuscript most probably copied in France on an Acre exemplar.
• The manuscripts fr20125 and Rennes share a very similar text.\textsuperscript{15} For the sections devoted to the Genesis and the Greek Vienna has a text close to fr20125 (and Rennes).
• The manuscripts of the Abridged version present a shortened text of the sections Thebes, Greeks and Amazons, Troy, Eneas, Assyrian kings and the first paragraphs of the first part devoted to Roman history.\textsuperscript{16}
• The manuscripts fr20125, Rennes and those of the Abridged family share a number of features and, most importantly, a significant error: an inversion in the correct sequence of a series of paragraphs.\textsuperscript{17} While this error implies that these manuscripts share a common exemplar, its absence from the manuscripts of the rest of the tradition means that they depend on different manuscript sources.
• The manuscripts of the Acre group were produced Outremer by the end of the 13th century. In spite of a great deal of common features, they feature some internal differentiation.
• With regard to the contents of the Histoire ancienne, the most conservative narrative sections across the tradition are those devoted to the Bible (Genesis up to Jacob) to Mesopotamian and Assyrian history (Orient I) and to Roman republican history (Rome II).\textsuperscript{18}
• The manuscript fr686 is an Italian manuscript but its linguistic features show that it was based on an exemplar produced in Acre.\textsuperscript{19}

\textsuperscript{15} The place of compilation of fr20125 is still under debate. See Zinelli 2016a, 110: “Le manuscrit fr. 20125, lui-même, a été considéré comme originaire de Terre sainte, soit (plus probablement), comme copié d’un modèle de Terre sainte”. See also Zinelli 2013.
\textsuperscript{16} This group was formerly identified by Jung (1996) as the $\beta$ group. Rachetta (forthcoming) clarifies the exact nature of the relationships between both the manuscripts of this group and fr20125 and this group and Vienna (see below).
\textsuperscript{17} Namely the inversion of Histoire ancienne §73–73: see Rachetta forthcoming.
\textsuperscript{18} See Meyer 1885, Jung 1996, and Trachslser 2013 for the subdivision in parts and the respective labels of the narrative sections of the Histoire ancienne.
\textsuperscript{19} Crucial for the relations between fr686 and the Acre tradition is Zinelli 2016a. With regard to the language of the manuscript see Zinelli 2016a: 113–4. As for the stemmatic “position” of fr686 within the Outremer tradition of the Histoire ancienne, see Zinelli 2016a, 114, n163: “Le nombre de cas où P10 [= fr686] coïncide en lacune avec DBLPa [= Acre group] pourrait suggérer que le manuscript soit à placer dans une branche commune à ces manuscrits au sein de laquelle il occuperait une position de supériorité (ce qui donnerait raison de ses quelques coincidences dans la bonne leçon avec $P [=fr20125]$). The results of the linguis-
• **Vienna** is a composite manuscript. According to Rachetta (forthcoming), this codex underwent different stages of editorial work, including abbreviation and amplification.20

As for the linguistic features, I have not defined a closed set of phenomena. Rather, I have established a twofold flexible criterion. In my approach, a linguistic feature is relevant if 1) it intersects two or more levels of analysis (e.g., graphemics and morpho-syntax; lexicon and syntax, and so on), and 2) if it can be described in terms of its distribution — i.e., through the description of the mutual relations that a certain (phono-morphologic, syntactic, lexical) element entertains with other elements within an environment or a set of environments (see below §2.1).21

In the following subsection I will focus on two linguistic features. In §2.1 I will consider the form and structure of hypothetical comparative clauses (pattern: Fr. *comme si* + (subject +) inflected verb = En. “as if” + (subject +) inflected verb). We will see how the description of this feature is related to how some graphemic-phonological material is transferred from one manuscript copy to another.22 In §2.2 the focus will be on the adverbial relativizer *ou ens* and on the alternative strategies of adverbial locative relativization adopted by the manuscripts. The examples in subsection §2.1 and §2.2 guide our understanding of the scribes’ perception of, and margins of tolerance to, marked linguistic features. An appropriate description of these features may help us in catching a glimpse of the existence or residual survival of discrete linguistic varieties within the Old French continuum. Perception and tolerance move between the scribes’ respect for the exemplar, regardless of the grammaticality of the transcribed form in their own

20. Rachetta’s findings may lead to a significant redefinition of the position of Vienna in the textual tradition of the Histoire ancienne. While waiting for Rachetta’s publication of her research, and given the fact that different sections of Vienna could need different linguistic treatment according to the editorial stage they mirror and the source(s) upon which they depend, I have decided not to include the linguistic data from this manuscript in the examples below.

21. I prefer to use the generic term ‘environment’ rather than textual string or sentence because the analysis can involve both.

22. While all the members of the team actively contribute to every aspect of the TVOF’s research agenda, the linguistic seam of the project is under my responsibility.
variety, and their conscious adoption of a marked feature to give the text a (stylistically) distinctive form.\textsuperscript{23}

\subsection*{2.1 Graphemics and Syntax}

In modern French (as in modern English), the structure of hypothetical-comparative clauses is based on the template: \textit{comme si} + nominal or pronominal subject + inflected verb (V) (which corresponds to English \textit{as if} + Subject + V):

(1)  
\textit{Il mangeait comme s'il n'avait rien bouffé depuis des journées}

In (1), both the hypothetical conjunction (\textit{si}) and the pronoun (\textit{il}) need to be expressed. Moreover, both are followed by a verb in the imperfect indicative. This structure is frequently attested in Old French, where the imperfect subjunctive mood is preferred to the indicative.\textsuperscript{24}

(2)  
\begin{quote}
\textit{Si se regarde et voit cheüe  
Sa mere al pié del pont arriere,  
Et jut pasmee en tel maniere  
Com s'ele fust cheüe morte.}
\end{quote}
\textit{(Graal, 620–3 [my emphasis])}

\textsuperscript{23}. By linguistic variety I mean a discrete unit placed within the Old French continuum. The continuum can be analyzed in varieties sharing a core of “partial similarities” and featuring a smaller but crucial number of “partial differences” (\textit{Weinreich} 1954, 395). In Weinreich’s terms, “the more pressing and more troublesome problem [with this approach] is [. . .] how to break down a continuum into discrete varieties. What criteria should be used for divisions of various kinds?” (\textit{Weinreich} 1954, 396). The problem about how to consider and reorder differences within the Old French continuum is a serious one. As stated above, I have adopted an approach whereby textual variants are always analyzed at more than one linguistic level. The common thread, however, is that any variant always involves a morpho-syntactic or syntactic issue. Hence the importance that syntax has in my work on both the \textit{Histoire ancienne} and Old French thirteenth-century prose more generally. “Grammar” is more likely to help us in identifying the salient trait of a linguistic variety or sub-variety. For a case whereby syntactic analysis gives us access to typologically different competing varieties in the same texts, see \textit{Ledgeway–Ventura} (forthcoming) on the syntax of negation in matrix and subordinate clauses in Old French.

\textsuperscript{24}. According to \textit{Moignet} 1988, 248: “l’imparfait de l’indicatif, exceptionnel au XIIe siècle, se développe au XIIIe”.

\textsuperscript{(1)} Il mangeait comme s’il n’avait rien bouffé depuis des journées

\textsuperscript{(2)} Si se regarde et voit cheüe  
Sa mere al pié del pont arriere,  
Et jut pasmee en tel maniere  
Com s’ele fust cheüe morte.
In the following examples from the *Histoire ancienne* I will discuss some cases where the comparative constructions do not entirely follow the *comme si + Pro + V* pattern. Contrary to modern French, in Old French it is possible to have elliptic constructions, such as:

(3)

Si par semble fiere beste,  
*com* les gens doie corre sus  
((Guillaume de Palerne, 5526–7)

Notice that in (3) the “if” conjunction and the subject pronoun are not expressed. Moreover this case favors the present over the imperfect subjunctive.

Example (4) below comes from the Rome II section of the *Histoire ancienne*.\(^{25}\) In *fr20125* reading, one of the elements of the hypothetical-comparative pattern is also missing, but the interpretation of which component of the hypothetical-comparative construction is not present is not as straightforward as it seems:\(^{26}\)

(4)

Entre les autres merveilles qui avenoient adonques en la cité de Rome et en la contree, avint l’an que la cités ot esté fundee ·cccc· et lxxx· ans tot droiturerement, une grans merveille qui mout fu perillouse a veir et orible a raconter et a dire. Quar en pluisors lius sorgoit sans des fontaines et corroit toz vermausaval les ruisseaus ausi *come ce fust aigue clere*.  
Et avec ceste merveille plovoit lais tos blans des nues et cheoit a grosses gotes si que la terre en estoit arosee ausi come ce fust de pluie. (*Histoire ancienne* §886.02 [my emphasis])\(^{27}\)

In this case, the interpretation of the graphemic and lexical material in the sequence *come ce fust aigue clere* has consequences for the comprehension of the syntax of the hypothetical-comparative structure and, as we will

---

25. The Rome II section is devoted to the history of the Roman Republic until the end of the civil war.


27. The source is Orosius 1500: “nam et plurimis locis scaturiens e fontibus cruor fluxit” (*Historiae adversus Paganos* IV 5).
see, of the whole passage. Is the form ce in (4) an expletive pronoun (En. “it”) or it could stand for a conjunction (i.e., ce = se “if”)? Given what we standardly know about non-V2 word order in embedded clauses, my first hypothesis would be to interpret ‘ce’ here as an expletive subject (inasmuch as null subjects are not licensed in non-V2 embedded clauses).

The essential elements of the dossier are as follows. First, the graphemic oscillation between ‹c› and ‹s› followed by a mid front vowel ‹e› is frequent in fr20125. Moreover, we can find it both in Picard and Outremer documents, including the Histoire ancienne’s manuscripts of the Acre group, and fr686, an Italian codex drawing on an Acre source (see Minervini 2010, Rochebouet 2015, Zinelli 2016a). On top of that, both ce/se as a pronoun or hypothetical conjunction (“if”) are attested in Old French.

Secondly, two other instances of ausi come ce + V are attested in the same Rome II section of the Histoire ancienne. The first occurs a few lines after the passage quoted above:

(5)

Et avec ceste merveille, plovoit lais tous blanc des nues et chiot a grosses gotes si que la terre en estoit arosee ausi come ce fust de pluie. (Histoire ancienne §886.02 [my emphasis])

The second occurrence can be found a bit later in the same section:

(6)

[et] que li plusisor furent si agrevé de famine que li vif mangoient les mors [et] devoroient ausi come ce fussen bestes sauvages derivees (Histoire ancienne §1058.02 [my emphasis]).

28. With regard to the word aigue, its occurrences in Outremer texts and interpretation, see Minervini 2010, and Zinelli 2016c.


30. A third case, another rain-based simile, can be found in the Rome II section: “ains traioient [et] lansoient [et] dune part [et] dautre si espessement a la uolee come ce fust pluie menuete ou prim tans dauril a la matinee” (§1124.4).

31. See Orosius 1500: “et de nubibus guttatim in speciem pluuiae lacte demisso, diri, ut ipsis uisum est, terram imbres irrigauernt” (Historiae adversus Paganos IV 5).
The guiding hypothesis is that these kinds of structures are grammatically marked. If so, the manuscripts could potentially show some degree of variation. The study of how manuscripts vary can help us describing the features and grasping the nature of the linguistic phenomenon more generally. To verify this, let us have a look at the variant reading: how do the manuscripts of the *Histoire ancienne* read in the case of examples (4) to (6)? Beginning with (4), while most the manuscripts of the main groups have the “standard” structure:

(7a) [= (4)]

there are some important exceptions, as in *Add19669*, a manuscript of the Abridged family.\(^{32}\)

<table>
<thead>
<tr>
<th><em>Histoire ancienne</em> §886.02 fr686 f. 304vb, fr9682 f. 239va, and Acre family (Add15268 f. 230rb)</th>
<th>Abridged family (Hague78D47 f. 136ra)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quar en pluisors lius sorgoit sans des fontaines et corroit toz vermaus aval les ruisseaus ausi <em>come ce fust aigue clere</em>.</td>
<td><em>com se ce fust aigue clere</em></td>
</tr>
</tbody>
</table>

(7b) [= (4)]

*Add19669* similarly reads with *fr20125* and differently from the manuscripts of its group in the case of example (5).\(^{33}\)

<table>
<thead>
<tr>
<th><em>Histoire ancienne</em> §886.02</th>
<th><em>Add19669</em> f. 172vb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quar en pluisors lius sorgoit sans des fontaines et corroit toz vermaus aval les ruisseaus ausi <em>come ce fust aigue clere</em>.</td>
<td><em>ens</em> com ce fust eue</td>
</tr>
</tbody>
</table>

\(^{32}\) *Add19669* is a thirteenth-century codex of the Abridged family that has been associated with other manuscripts all compiled in the Picard area. Namely: Aylsham Blickling Hall, MS 6931; Lisbon, BN, Illum. 132; Paris, BNF, f. fr. 17177 (= *fr17177*); Pommersfelden, SW-SS 295; Den Haag KB 78 D 47 (= *Hague78D47*). As for the illumination cycle of this group, see *Rodríguez Porto* 2013.

\(^{33}\) Paragraphs §886 and §1058 are missing from *fr17177*. But in §1124 is present: “*com se ce fust pluie menuete*” (c. 176va).
On the other hand, with regard to example (6), the family has the standard structure:

(7d) [= (6)]

| ausi [com]sece fussent bestes sauuages des|uees (Add19669 207ra, Hague78D47 c. 166rb [my emphasis]) |

If we consider again examples (4) to (6) from the broader perspective of Old French textual and literary tradition, examples like the following from the Anglo-Norman biblical translation known as the Quatre livres des Rois, show that the form ce in hypothetical-comparative clauses was interpreted as pronouns:34

(8)

Absalon out fait un cunvivie si cume çó fust le cunvívie le réi (Quatre livres des rois, 82 [my emphasis])35

This example may not exclude other interpretations of ce. First, notice the opposite order ‘si comme’, rather than ‘comme si’. Secondly, the oscillation ‹c›/‹s› is not an uncommon graphemic trait in early Picard manuscripts. There is evidence that in the Histoire ancienne’s tradition ce could stand for se (“if”) in hypothetical clauses, albeit not in comparative structures. See the following unambiguous case in Add19669:

(9)

Mais sachiez qu[ue] | ce ne fust la grant force de hercules que eles [= the Amazons] | eussent mal mene les grezois (Add19669, c. 76ra [my emphasis])36

34. Cf. TL 2,83, 16 and most notably 23, where Tobler considers ce as the subject of athematic verbs (estre and so-called impersonal verbs; see also Buridant 2000, §108).

35. “Fecitque Absalom convivium quasi convivium regis” (II Sam. 13,27). In Quatre livres de rois this construction is well attested: I have counted no less than nine occurrences.

36. Comparing Add19669 with Hague78D47, we observe that the latter rephrases (9): “Mes sachiez ne fust la | gfranf force hercules q[ue] eles eussent mal mene |
In both fr20125 and the Acre family the sequence is se ne: “... se ne fust la grant force ...” (fr20125 §509.08, Add15268 f. 104ra, D562 c. 87vb ff. [my emphasis])

Secondly, I wonder whether this is the situation mirrored by fr686, an Italian manuscript copied from an Acre exemplar. This oscillation could have produced a polymorphism whereby the scribes could have had //ce ~ se// as an alternative for “it” and “if”. As regards (5), fr686 reads:

(10)
aussi come se fust de pluie (fr686 c. 304vb [my emphasis])

The hesitation as to whether the scribe of fr686 could have interpreted se as a hypothetical conjunction rather than as a pronoun written with an ‹s› is confirmed by how fr686 reads for (6):

(11)
aussi co[m] | se fussent bestes sauuages desuees (fr686 c. 366va [my emphasis])

The copyist could have tolerated the ambiguity in se (= ‘if’ / ‘it’?) on two counts: the already noted ‹c›/‹s› graphemic oscillation for [s], and the similarity between this French construct with the Italian equivalent come se + inflected verb structure.

We can now try to represent the presence of the (4), (5), and (6), where the hypothetical-comparative in asyndeton/polisyndeton (with or without the conjunction), in tabula form representing the textual tradition of the Histoire ancienne:

les grezois” (Hague78D47 c. 62rb).

37. See Zinelli 2016a: 104: “Pour le consonantisme, l’emploi frequent de s pour c (sil, siaus, renonceront etc. [. . .]), correspond souvent aux habitudes des scribes italiens au travail sur des textes français”.

38. In spite a minor difference in the predicative clause, both fr9682 (“aussi come se ce fust pluie”, fr9682 c. 239va) and the Acre tradition has the standard structure (“aussi come se ce fust de pluie”, Add. 15268 c. 230va).


40. See the following examples from Dante’s Vita Nuova: “E che io dica di lui come se fosse corpo, ancora si come se fosse uomo, appare tre cose chi dico di lui” (VN XXV); “e detto che molti accidenti parlano, si come se fossero sustanzie e uomini” (VN XXV [my emphasis]).
Table (12a) indicates that the unmarked structure (\textit{com me se + ce + V}) is present in all the manuscripts of the tradition. Table (12b) shows the uneven distribution of the marked structure (\textit{com me se/ce + V}) across the tradition. Traces of the marked syntactic structure are present just in some of the manuscripts of the \textit{Histoire ancienne}. For reasons of economy, we can therefore assume that it was the marked structure the most likely to be levelled down in the tradition and not the reverse.\footnote{For an application of the sociolinguistic and contact linguistics notion of leveling in medieval textual traditions, see Zinelli 2016a and Zinelli 2016b.} This means that the marked structure was not an idiosyncratic linguistic option singling out a single manuscript or a group of manuscripts. Rather it is a conservative trait shared by a number of manuscripts pertaining to different families of the tradition. The standard structure represents the innovative and polygenetic tendency of the scribes to level down markedness. The marked structure represents a salient feature of the oldest stages in the textual tradition of the \textit{Histoire ancienne}.

The two-layered tables above (12a)–(12b) schematize this state of affairs. Table (12b) (pattern: \textit{come se/ce + V}) represents the distribution of the marked and conservative trait. The unmarked/standardised ‘layer’ on table (12a) (pattern: \textit{come se ce + V}) represents the innovative and polygenetic tendency of the tradition to level down the marked trait. On the one hand, the tolerance of some of the manuscripts of the \textit{Histoire ancienne}, namely \textit{fr20125} and \textit{fr686} (and occasionally in \textit{Add19669}), to this form singles out the shape of the ‘version’ of the work common to and passed down by these manuscripts. On the other hand, the manuscripts of the Acre family and particularly of the manuscripts of the Abriged family show a lower degree of tolerance to the elliptic hypothetical-comparative structure that, by the end of the 13th century, must have been considered a diachronically marked (‘archaic’) feature.\footnote{In my research on the language of the \textit{Histoire ancienne}, I could verify the same ‘behavioural’ trends in the textual tradition of this work. Two clear examples}
3. Conclusion

Examples in section §2.1 illustrate how textual variation is linked with specific linguistic features at graphemico-syntactic level. This connection is based on the comparison of a sample of manuscripts chosen on the basis of the digital edition of the complete text of the Histoire ancienne (based on fr20125), and the availability of a tool, Alignment, a map of the contents of the Histoire ancienne in its rich manuscript tradition. Alignment has proven an invaluable instrument in understanding the relations between the manuscripts of the Histoire ancienne. At the same time, we would not have realized the ‘competition-behavior’ in language among the manuscripts without Alignment and the systematic study of the textual tradition (see also Ledgeway–Ventura forthcoming).

From a linguistic perspective, the first redaction of the Histoire ancienne as it appears in fr20125 and — to a different extent — in fr686 Add19669 and Rennes, reflects a peculiar form of the text different from that passed down by the rest of the textual tradition. For reasons of economy, it is more sensible to suppose that fr20125 (and fr686) had a conservative approach to the language of its sources than the other manuscripts. The rest of the tradition tended to level down those linguistic features that might have been perceived as ‘marked’, albeit in different measures. Now, the manuscript fr20125, those of the Acre family and of some the manuscripts of the Abridged family were all compiled at the end of the 13th century.

The linguistic refashioning of the first redaction of the Histoire ancienne could have started relatively early, but its traces became visible only at a later stage, during the last decades of the century, when features like those exemplified in §2.1 might have begun to feel structurally marked and stylistically old-fashioned or archaic. We also know that the editorial work witnessed in the manuscripts of the Abridged family is anterior to 1260 (Rachetta 2018). The behavior of manuscripts such as Add19699, where some of the marked features are still present, indicates that the linguistic editing was progressively done after the archetype of the Abridged family was composed, hence after 1260. The almost complete absence of the marked features described above in the manuscripts of the Acre family shows how the margin of tolerance for certain linguistic characteristics shrank over the last two decades of the 13th century. Yet, the conservation of these features in late thirteen-century or early fourteenth-century man-

---

are the syntax of the relative clauses and the syntax of negation: see Ventura, forthcoming and Ledgeway and Ventura, forthcoming.
uscripts pertaining to different families of the *Histoire ancienne*’s textual tradition, such as fr20125 and fr686, shows that we are not facing “singular” innovations of an idiosyncratic scribe. It means that these features were most probably in place at a very early stage in the textual transmission.  

King’s College London

**Works Cited**


43. At question here is not what textual scholarship traditionally labels *lectiones singulares* (singular variant readings). Important as they are linguistically and culturally, this kind of variant reading is not taken into account for the purposes of textual reconstruction.


