



HAL
open science

The evolution of the syntax of the subject in French and factors of variation

Sophie Prévost

► **To cite this version:**

Sophie Prévost. The evolution of the syntax of the subject in French and factors of variation. Janice Caruthers; Mairi McLaughlin; Olivia Walsh. Historical and sociolinguistic approaches to French, Oxford University Press, pp.72-96, 2024, 9780192894366. <hal-04681558>

HAL Id: hal-04681558

<https://hal.science/hal-04681558v1>

Submitted on 29 Aug 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

4 The evolution of the syntax of the subject in French and factors of variation

Sophie Prévost

Lattice, CNRS/Université ENS-PSL/Université Sorbonne Nouvelle

4.1 Introduction

The notion of variation can be approached from various angles: diachronic, diatopic, diastratic, diaphasic, and diamesic. For the researcher who works on linguistic change, the diachronic dimension comes first. However, disregarding the other aspects would result in drawing an incomplete picture of the phenomenon under study: a change does not spread equally in linguistic structures, communicative situations or groups of speakers. Considering these different aspects makes it possible to fully account for the modalities of change, as regards differentiated chronologies as well as modalities of propagation of new variants. Identifying the pioneering loci of change may also provide some answers to the thorny issue of the causes and motivations for the changes. However, studying language states prior to the seventeenth century makes access to diastratic, diaphasic, and diamesic dimensions very difficult due to the lack of relevant data. On the other hand, we can consider other criteria of variation such as the domain and/or the genre of the texts as well as their form (verse or prose), and the enunciative situation (direct speech versus narrative), which highlight noticeable differences regarding the embedding of linguistic structures in language and their evolution. This study is devoted to two major changes that took place in French between the twelfth and the seventeenth centuries: the increase in overt subjects and their growing anteposition to the verb. Although these two changes are partly related from a linguistic point of view, a close examination of the data shows that they are not equally sensitive to the following criteria of variation: time, dialect, domain, form, enunciative situation (direct speech/narrative), and verbal person. The study aims to account for both changes while attempting to highlight the specific influence of each of these different factors and to explain this contrasted situation.

4.2 The historical linguist faced with bad data

4.2.1 When? How? Why?

When studying any linguistic change, the historical linguist has to answer three fundamental, and partially related, questions: When? How? Why? The first question aims to identify not only the points of departure and of arrival, but also the intermediate stages of a change. Studying changes that took place in remote past states of the language offers at least one advantage: the passage of time allows us to determine with certainty the endpoint of a change, based on the stabilization of the variants (which does not preclude the possibility of a new change). Spotting the starting point of a change is less straightforward (regardless of the period under scrutiny), whether the change consists in the emergence of a new variant which gradually wins out over an already existing one, or in a change in the equilibrium between two (or more) existing variants. In both cases, one has to decide when the change really started, which does not equate with the mere advent of an innovation (whatever its form): it has to propagate through contexts (Andersen 2001). The quantitative aspect is certainly a key criterion (though it is not possible to give an absolute threshold), but it needs to be refined with a qualitative approach, especially as regards what exactly is meant by ‘context’.

The ‘how’ question is actually a twofold one, as it concerns both the mechanisms involved in a change and the modalities of its propagation – an aspect which is related to the previous question. All studies based on corpora show indeed that a change does not spread evenly in the different loci of a language. Finally the ‘why’ issue probably remains the most challenging one, as the identification of causes and motivations of a change remains to a large

extent speculative, all the more so the further back into the past that we step. However, the answers to the preceding questions may provide some enlightening insights.

4.2.2 Variation and diachronic syntax

Sociolinguistic studies have enriched the ground of diachronic studies by emphasizing the social dimension of language and the fundamental role of variation.¹ These factors had in fact been already pointed out a few decades earlier, especially by Antoine Meillet (e.g. 1916). Jakobson also highlighted the prime role of variation in change: ‘Pendant un certain temps, le point de départ et le point d’aboutissement de la mutation se trouvent coexister sous la forme de deux couches stylistiques différentes [...] Un changement est donc, à ses débuts, un fait synchronique’ (Jakobson 1963: 37).

There is now a wide consensus that, except in the domain of the lexicon, a change always results from a variation.² The first task consists in identifying the nature of the variation as well as the variants at stake. In the two cases under study here – the growing use of overt subjects and their increasing preverbal position – we are not dealing with the emergence of new variants. Null subjects and overt subjects on the one hand, and preverbal subjects and postverbal subjects on the other hand have always been attested and the change thus consisted in the evolution of their frequencies, namely a sharp decrease in null and postverbal subjects, and an increase in overt and preverbal subjects.

However, one may wonder how appropriate it is to consider that these are cases of variation. This question falls within the more general and long-debated issue regarding the relevance of addressing syntactic changes from a socio-linguistic perspective. The fact that Labovian socio-linguistic methodology was developed for phonology means that its application to syntax is far from self-evident. As Ayres-Bennett (2004: 10) recalls, ‘the concept of the linguistic variable has been successfully applied to the phonological variation because of the arbitrariness of the linguistic sign [...] In the case of syntax the problem of determining identity of meaning for all variants of the variable is more acute. Gadet (1997: 11) suggests that this necessitates a looser interpretation of functional equivalence [...]’. To identify the exact ‘meaning’ of two variants, and their equivalence or not, is all the more difficult when it comes to past language states, for which we do not have any native speaker’s intuition. However, in the present case, a close examination of the contexts shows that, first, the variants were partially constrained by semantics (a null subject was used only if the referent was cognitively active), syntax (the verb-second constraint) and pragmatics (informational weight of the subject), and second – and this is in part correlated to the previous point – that there were semantic and/or pragmatic nuances opposing overt subjects to null subjects, and postverbal subjects to preverbal ones. For reasons given in section 4.2.3, it can be claimed that the variation was not related to any social dimension. Nor was it associated with any stylistic effect, at least during the early stages: however we do find such effects in certain uses of postverbal subjects in Modern French. In the light of what precedes, the notion of ‘variant’ may be questioned, if taken in a strict Labovian sense. However, following Gadet (1997: 11), I will adopt ‘a looser interpretation of functional equivalence’, considering that null and overt subjects, be they preverbal or postverbal, when associated with a verb, are all variants of the same meta-function denoting the predicative relation between a verb and its subject.

¹As underlined by Ayres-Bennett (1994: 3), the interest in socio-historical studies of French is relatively recent (starting in the 1990s: Ayres-Bennett (1996); Lodge (1993; 2004); Posner (1997), for instance) in comparison with studies on English (Romaine 1982). Ayres-Bennett was a remarkable pioneer in this domain, with landmark studies of seventeenth-century French language.

² See Prévost and Dufresne (2020: 23-7) for a discussion.

4.2.3 Internal and external contexts; parameters of variation

Changes propagate through time but also through different types of contexts, whether internal (linguistic structures) or external (situations of communication and groups of speakers), and they do so at different paces; moreover, they do not necessarily spread in all contexts. To ignore the contexts of variation and of diffusion of a change gives a very impoverished picture of the phenomenon and it also prevents us from detecting some of the motivations for the changes, as the pioneering contexts may sometimes offer some clues as to why a change started out. The different contexts can be reinterpreted in terms of factors of variation, and from the point of view of the analyst, in terms of parameters of variation.

In contrast with the linguist studying present-day changes, the historical linguist has no direct knowledge of the way in which changes propagated through groups of speakers and through situations of communication. First, we do not have any recording of the spoken language before the late nineteenth century, and therefore we can only rely on written data. Second, there are no documents illustrating informal or semi-literate usages before the seventeenth century (and even at that time, there are few of them). Medieval documents all stem from literate writers, at a time when only a very small percentage of the population could read and write. Moreover we do not have any reliable testimony of non-standard usages, as was the case from the seventeenth century on, thanks to the comments of the *remarqueurs* and grammarians. There might have existed low-literate documents in Medieval French, but if it were the case, they have unfortunately not reached us. We are also ignorant of the age and the gender of authors, and thus the influence that these parameters may have had on their writings. Until the seventeenth century, many documents were indeed anonymous, and even when the author's name is known, it is often difficult to establish the exact dating of the document and/or the date of birth of the author. In view of the expected methodology of a true sociolinguistic approach and of the above-mentioned limitations, the handling of past variations, and especially of Medieval ones, is undoubtedly problematic since we are unable to take into account their diastratic, diaphasic, and diamesic dimensions.

However, we can consider other types of contexts. Although they do not have any social value, they allow us to observe how the changes progressively spread through written data according to external parameters. These are the domains of the texts (literary, historical, argumentative, juridical/legal, hagiographical/religious, didactic, and epistolary),³ and their form (verse or prose), the hypothesis being that texts in prose were both less constrained and possibly more innovative, especially as prose developed later than verse. We can also consider, when relevant, some more specific features of the texts, such as the register. Although medieval texts pertain to formal and high-level registers, they may display some more spontaneous passages, in direct speech. This point leads us to address the critical question of spoken language. The fact that we do not have any access to spoken language of past language states deprives us of important information relating to the material dimension of the language (prosody, phonetics ...) but also of access to informal registers. It is not a question of strictly equating spoken language with low-register usages, but spoken language may display less formal usages, which are not to be found in written data of this period, and it may, for this very reason, constitute a pioneering context for some changes, as can be observed in Modern French. In the absence of access to spoken language, we can consider direct speech, with, needless to say, due caution. I certainly do not claim that direct speech can be equated, whatever the period, with spoken language, since it may display over-corrected usages or, on the contrary, exaggerated and caricatural features. It may, however, be considered as a pale and/or imperfect

³ These domains were identified by the *Base de Français Médiéval*'s team (<http://corptef.ens-lyon.fr/spip.php?article62>) and stand as a reference among historical linguists working on Medieval French. Accessed in November 2020.

reflection of spoken language and provide some interesting information, especially if it displays the first signs of some changes, in comparison with narrative passages.

Let us now turn to internal contexts. Just as changes do not spread at the same pace through external contexts, they do not propagate evenly through linguistic structures. As regards the evolution of the syntax of the subject, the following factors turn out to be relevant: the type of clause (main declarative versus subordinate clauses), the presence of another argument, especially the object, and the verbal person (the latter factor being partially correlated to the opposition between direct speech and narrative, as we will see in section 4.4.4).

4.2.4 Collecting the data

Collecting the data raises two difficulties, especially for past language states: the gathering of sufficient data illustrating the variables, and the composition of a balanced corpus. The first largely depends on the phenomenon under scrutiny. The evolution of the expression and of the position of the subject is not a rare phenomenon and therefore data is not lacking: all texts display an abundance of null and overt subjects, as well as preverbal and postverbal ones, albeit in varying proportions (which is precisely the change to be studied). The handling of internal parameters is not a problem either: all texts comprise main and subordinate clauses, transitive and intransitive structures, and, though less systematically, a diversity in verbal persons. Although all texts do not display stretches of direct speech, sufficient can be collected data to compare them with those in narrative passages.

Things are less straightforward when it comes to external parameters, namely the date, domain and form of the texts. The further back into the past we go, the scarcer the texts are, and it is difficult to gather more than a few of them till the late eleventh century. They are so scarce, in fact, that we do not have to wonder which ones to select, but a balanced corpus cannot be built according to date, domain and form. The situation improves from the twelfth century on, since many more texts from that period have reached us, but, till the thirteenth century, they are mainly religious, in verse, and Anglo-Norman, making it often impossible to separate out the different parameters. It is only from the thirteenth century on that the available texts are sufficiently diversified for the parameters to be handled separately, though certainly not as rigorously as for modern language states. The increase in texts over the centuries raises the question of how to build a corpus that is sufficiently representative from a quantitative and qualitative point of view, while remaining manageable as regards data collection and processing, especially when the changes studied are instantiated by a large amount of data.

Two last points deserve to be mentioned as regards the value of some parameters across centuries. First, as may have been noticed, I have not explicitly mentioned dialectal variation (which is acknowledged to have no longer operated, generally speaking, beyond the late fifteenth century). Syntactic variation was indeed much less sensitive to the diatopic dimension than morphological variation, particularly as regards the evolution of the syntax of the subject (see Prévost and Marchello-Nizia 2020: 1059). A close examination of my data corroborates this statement, and this parameter will thus be alluded to only when relevant. Second, as regards the form of the texts, it must be pointed out that the opposition between prose and verse has evolved over the centuries. With a few exceptions, till the late twelfth century, all texts were written in verse: prose was an exception, and verse was the unmarked form. The thirteenth century offers a more balanced picture, with a growing use of prose, especially in didactic and historical texts. However, verse remained very much alive, especially in literary texts, even though the use of prose increased in this domain as well. From the fourteenth century on, the use of verse started to lose ground, and progressively became restricted to specific literary genres: poetry, songs, etc. In a word, while verse was the default form till the thirteenth century, it gradually became a stylistically marked form.

The paucity of documents, especially for the most remote past language states, the need to rely only on written data, the difficulty of separating out some parameters of variation – given these limitations, we may wonder how reliable and representative the sources are, and therefore how sound and valid the descriptions of changes and their interpretations are. Such a situation led Labov, a few decades ago, to label historical data as ‘bad’, in words which have become a sort of aphorism (‘the great art of the historical linguist is to make the best of this bad data’). This remark deserves to be placed in a larger context, however, in order to clarify exactly what is meant by ‘bad’:

The fundamental methodological fact that historical linguists have to face is that they have no control over their data. Texts are produced by a series of historical accidents; amateurs may complain about this predicament, but the sophisticated historian is grateful that anything has survived at all. The great art of the historical linguist is to make the best of this bad data – ‘bad’ in the sense that it may be fragmentary, corrupted, or many times removed from the actual productions of native speakers. (Labov 1972: 100)

As Romaine (1988: 1454) pointed out, historical data are ‘bad’ only if we try to put them on an equal footing with contemporary data and especially with spoken ones. There are inherent limitations in studying them, and it must be admitted that they constitute a specific material, valid on their own, but quite different from modern states of the language. We therefore have, legitimately, to adapt our methodologies, and to accept that our claims about past linguistic structures, and their evolution, are only an interpretation of the available data. The more substantial and balanced the corpus under study is, the more our interpretation may come close to what past languages really looked like, and when and how evolutions proceeded.

It would certainly be exaggerated, and even false, to claim that the present study falls within a socio-historical approach. More modestly, it aims to account for two major – and related – changes which have occurred in French, while taking into account certain parameters of variation that provide a more detailed picture of the evolution, as regards the ‘when’ and ‘how’ questions, and even, to some extent and with due caution, that make it possible to suggest some possible motivations for the changes.

4.3 A quick overview of the changes under study

Modern French is a fairly strict SVO language (at least in main declarative and subordinate clauses, on which I am focusing here), in which the expression of the subject is obligatory (with a few exceptions, among which contexts of close coordination or juxtaposition between verbs). These characteristics sharply distinguish Modern French from the earliest stages of the language.

Old French had inherited from Latin the possibility of not expressing the subject when its referent could be unambiguously identified. As a result, until the end of the twelfth century null subjects were very frequent and they even prevailed in main declaratives. A first major turning point took place at the beginning of the thirteenth century, and a second one occurred in the mid-sixteenth century, with frequencies of overt subject expression rising to more than 90%.⁴ Old French was moreover characterized by the existence of a flexible word order (as regards the grammatical functions), allowing for preverbal objects, postverbal subjects, and for all six permutations of S, V, and O. The evolution towards an SVO order spanned several centuries. The earliest change concerned nominal objects (the preverbal position of pronominal objects, inherited from Latin, was settled at the very beginning of French, with a few exceptions; see Marchello-Nizia 2020). While in Early Old French (until 1100) nominal objects

⁴ See Prévost and Marchello-Nizia (2020: 1055-74).

prevailingly occupied a preverbal position, as was the case in Latin, OV order had fallen to 35% a century later (*Roland*, c. 1100), then to less than 10% in *Graal* and to only 5% in the early fifteenth century (Marchello-Nizia 2020).

In Modern French the position of the subject is more flexible than that of the object, at least in main declaratives and in subordinate clauses. However, postverbal subjects have receded considerably since Old French, along paths which differ according to their nature, nominal or pronominal. Postverbal nominal subjects did not decrease in a linear fashion, as will be seen in section 4.5, but the mid-sixteenth century marks a turning point: preverbal subject rates have remained over 80% ever since (Prévost and Marchello-Nizia 2020: 1080). The evolution of pronominal subjects (personal pronouns) offers a quite different picture. As they already largely prevailed at the very beginning of French, they did not undergo a very significant increase over the centuries. From more than 80 % in most texts, they had reached more than 90% by the early fourteenth century (Prévost and Marchello-Nizia 2020: 1080).

The reasons why null subjects were progressively lost in French (contrary to most other Romance languages) and why French evolved towards a strict word order is a long-debated, twofold question. Traditional explanations relate the two changes to some other morphological and syntactic changes.⁵ Old French possessed both a rich verbal morphology, which differentiated all the verbal persons, and a nominal declension, inherited from Latin (albeit reduced to nominative and oblique cases), which made it possible to discriminate the main grammatical functions. Old French was also characterized by the existence of a loose verb-second constraint, resulting in the postverbal positioning of the subject when the preverbal position was occupied by another element. This constraint has also been put forth as an explanation for null subjects, based on the claim that null subjects were actually postverbal pronominal subjects (inverted because of the occupation of the preverbal position by another element), which had been omitted. Relying on the existence of some sort of inverse dependency between a rich morphology (verbal endings and nominal declension) and a relative syntactic freedom (null subjects and flexible position of the arguments), it has been claimed that the morphological erosion of inflectional endings expressing subject-verb agreement on the one hand, and the disappearance of the nominal declension on the other entailed the generalization of overt subjects and the fixing of word order, especially the subject in preverbal position, henceforth compatible with the presence of another preverbal element, due to the loss of the verb-second constraint.

However, both the morphological and syntactic explanations have been called into question. As regards the relation between the loss of verbal morphology and the growing overt expression of subjects, the diachronic scenario was questioned by Schøsler (2002) and Roberts (2014) on the grounds of an apparent time lag between the two changes, though they make opposite claims about the temporal sequence of changes. Schøsler assumed that null subjects were largely lost at the beginning of the Medieval French period, while verbal inflection continued to undergo phonological restructuring for several centuries more, whereas Roberts rejected a direct connection between the two phenomena because of an apparent time lag. Based on an extrapolation from written to oral data given in Foulet (1935), and the completion of the loss of null subjects only by the sixteenth century (for a detailed discussion, see Prévost 2018; Simonenko, Crabbé, and Prévost 2019),⁶ he dates the total loss of rich agreement to the twelfth

⁵ See among others Foulet (1930 [1919]: 37-8; 446-61); Skårup (1975: ch. IV); Adams (1987); Vance (1997: 2; ch. 2; 3; 5).

⁶ It should be mentioned that there are some biases which make it difficult to assert the possible connections. There is indeed a methodological flaw in speculating on the pronunciation of morphological endings while relying on written data as concerns the frequency of overt subjects: it may be the case that subjects were more often expressed in spoken language. The time lag between verbal syncretization and the generalization of overt subjects may thus have been shorter than claimed by some linguists. Moreover, we may wonder what is meant exactly by 'loss', and whether it amounts to a mere quantitative approach, with the determination of a threshold.

century. As for the relation between declension and the identification of functions, first, even in a stage in which the decay of the case system was already advanced, there was still a flexible word order; and second, as Schøsler (1984) pointed out, there existed, from Latin onwards, other factors (among which valency and animate character) that enabled subjects and objects to be distinguished. Finally, the influence of the verb-second constraint and the consequences of its loss are challenged by the fact that it was not a strict rule in Old French (we find verbs in initial position, or, on the contrary, in third position, or even further on in the clause).

Although none of these factors should be dismissed, they are not sufficient to fully account for the existence of null subjects and for the variation in word order in Medieval French, nor for the loss of the former and the fixing of word order afterwards. This strictly morpho-syntactic approach was supplemented by pragmatic and informational explanations. First, the increase in overt subjects, more specifically in personal pronouns, has been related to a need for expressiveness (or emphasis). This hypothesis was put forward early (Foulet 1930 [1919]), and taken up and further developed by Detges (2003) (see Prévost 2018 and Prévost and Marchello-Nizia 2020 for a detailed account). The increasing use of pronominal subjects for the sake of expressiveness allegedly brought about some sort of rhetorical devaluation, which in turn led to a more widespread – and thus pragmatically less marked – use of overt pronouns. We will see in the following sections that this hypothesis is supported by a close examination of the data. Second, the flexibility of word order has been accounted for by an informational principle, formulated either in terms of Topic-Comment (Lambrecht 1994) or in terms of Theme-Rheme (Firbas 1992). As the subject was a privileged topic or theme (especially the personal pronoun), it often appeared in preverbal position, and this position was therefore progressively reinterpreted as the position of the subject. French, it is argued, went from a word order based on an informational principle to a grammatical one, based on the ordering of the syntactic functions (Vennemann 1976; Combettes 1988).

As is often the case, it is most likely that different factors – morphological, syntactic and pragmatic – jointly contributed to both changes.

4.4 The progressive obligatorification of overt subjects

As was recalled in section 4.3, Old French had inherited from Latin the possibility of not expressing the subject when its referent could be unambiguously identified. However, it would be an exaggeration to consider that OF was a true null-subject language (NSL). It is more appropriate to analyze it as a partial NSL. First, overt subjects prevailed in subordinate clauses right from the very beginning of French, and second, none of the supposed features of a NSL are observed: we find referential pronouns even in non-emphatic contexts and in the absence of any ambiguity, and expletive pronouns even in the earliest texts (ninth to eleventh centuries), albeit very rarely. While null subjects were very frequent and even prevailed in main declaratives till the end of the twelfth, a first major turning point took place at the beginning of the thirteenth century: subject expression rose from 47% in *Passion* (c. 1000) to 79% three centuries later (*Graal*, c. 1225). A second turn occurred in the mid-sixteenth century: from this time on, subject expression overwhelmingly dominated, with frequencies rising up to those we find in Modern French (more than 90%), though there remained till the end of the seventeenth century some constructions that are no longer acceptable (see Prévost and Marchello-Nizia 2020: 1055-79).

This increase was in favour of pronominal subjects (personal pronouns). The data clearly show an inverse dependency between the decrease in null subjects and the rise of pronominal subjects, which is no surprise if one considers the proximity of null subjects and personal pronouns on accessibility hierarchies (Ariel 1988), both of them signalling a high degree of cognitive activation. The other types of subjects remained more or less constant over the centuries. Therefore when dealing with the evolution of subject expression it would

probably be more appropriate to consider only personal pronouns. However, as studies may noticeably vary on this point, both rates are indicated in Table 4.1, which provides a quantitative overview of the evolution of overt subjects (the data stem from the extensive study which was conducted in Prévost and Marchello-Nizia 2020).⁷

Table 4.1 Rates of overt subjects (= all subjects: S) and overt pronominal subjects (= personal pronouns: Spp) in main declarative vs. subordinate clauses from the 11th to the 16th century⁸

<i>Texts</i>	<i>Domain</i>	<i>Rate of S</i>	<i>Rate of S in Declaratives</i>	<i>Rate of S in Subordinates</i>	<i>Rate of Spp in Declaratives</i>	<i>Rate of Spp in Subordinates</i>
<i>Passion</i> (c. 1000)	relig.	47%	43% (162/362)	60% (71/119)	16% (37/237)	27% (18/66)
<i>StLegier</i> (c. 1000)	hagiog.	43%	42% (77/182)	48% (37/77)	15% (19/124)	33% (20/60)
<i>StAlexis</i> (c. 1050)	hagiog.	48%	41% (212/514)	65% (147/227)	11% (36/338)	42% (58/138)
<i>Roland</i> (c. 1100)	literary	51%	51% (345/675)	67% (136/205)	16% (63/393)	41% (48/117)
Lapidaire (mid-12 th)	didact.	81%	74% (325/440)	92% (270/295)	55% (138/253)	74% (77/102)
<i>EneasI</i> (c. 1155)	literary	47%	38% (258/673)	69% (205/299)	13% (60/475)	47% (83/177)
<i>BeroulTristan</i> (late 12 th)	literary	55%	44% (283/639)	74% (220/298)	16% (70/426)	54% (92/170)
<i>TroyesYvain</i> (1177-81)	literary	59%	44% (203/464)	78% (355/455)	19% (61/322)	64% (19/279)
ClariConstantinople (after 1205)	historic.	75%	52% (262/504)	96% (486/505)	24% (75/318)	92% (214/232)
Aucassin (late 12 th -13 th)	literary	65%	47% (281/593)	90% (295/329)	31% (141/453)	79% (130/164)
CharteChièvres (1194)	legal	90%	89% (42/47)	90% (82/91)	55% (6/11)	79% (34/43)
CharteArras (1224)	legal	95%	100% (4/4)	93% (14/15)	-	86% (6/7)
Graal (c. 1225)	literary	79%	60% (300/497)	96% (470/490)	37% (117/314)	93% (264/285)
<i>RenartDole</i> (early 13 th)	literary	68%	61% (325/535)	86% (332/388)	36% (120/330)	74% (163/219)
Beaumanoir (1283)	didact.	88%	66% (183/278)	96% (693/720)	44% (75/170)	91% (258/285)
Joinville (1309)	historic.	82%	68% (286/419)	96% (510/533)	54% (157/290)	93% (300/323)
<i>MachautFortune</i> (1341)	literary	63%	57% (228/400)	69% (390/566)	40% (114/286)	48% (160/336)
Mesnagier (1393)	didact.	65%	67% (225/334)	85% (398/470)	43% (78/183)	76% (224/296)
<i>Griseldis</i> (1395)	literary	56%	45% (193/428)	70% (315/448)	31% (106/341)	52% (147/280)
Manières (1396, 1399)	didact.	76%	85% (332/391)	85% (306/358)	81% (253/311)	78% (190/243)
QuinzeJoies (c. 1400)	literary	79%	66% (283/426)	90% (457/512)	55% (173/316)	82% (255/310)
GersonSermon (1402)	relig.	83%	72% (240/331)	95% (342/361)	57% (122/214)	88% (142/161)
PizanCité (1404-5)	didact.	70%	57% (190/336)	83% (497/599)	31% (65/211)	63% (177/279)
<i>Pathelin</i> (1456-69)	literary	85%	89% (402/453)	93% (242/260)	86% (309/360)	90% (163/181)
CentNouvelles (1456-1467)	literary	78%	64% (234/365)	87% (508/581)	50% (133/264)	77% (240/313)
<i>ArchieBaignollet</i> (1468)	literary	81%	87% (162/186)	90% (87/97)	83% (118/142)	82% (46/56)
Commynes (1490-1505)	historic.	78%	64% (297/469)	91% (482/531)	32% (81/253)	79% (182/231)
JehanParis (1494)	literary	73%	59% (250/424)	87% (423/488)	42% (124/298)	74% (182/247)
Vigneulles (1515)	literary	75%	57% (260/458)	92% (423/460)	42% (146/344)	87% (252/289)
CalvinLettres (1549)	epistol.	95 %	95% (318/336)	97% (617/638)	93% (240/258)	95% (404/425)
Rate of S: less than 50%			Rate of S: between 50% and 80%		Rate of S: more than 80%	

⁷ Most rates are based on text samples containing 1000 finite verbs. Only the following texts were analyzed exhaustively: *Passion*, *SaintLégier*, *SaintAlexis*, *Lapidaire*, *Charters*, *GersonSermon*, *ArchieBaignollet*.

⁸ The rates of overt S and of overt Spp are calculated out of: 'overt S + null S' and 'overt Spp + null S'.
Italics: texts in verse; roman: texts in prose.

Table 4.1 highlights substantial discrepancies between frequencies, especially until the late thirteenth century, which are partly influenced by different factors or contexts, either external or internal, which are successively addressed in the following sections.

4.4.1 The form and domain of texts

The data in Table 4.1 clearly show that texts in prose are more advanced than texts in verse as regards subject expression. In the mid-twelfth century the frequency of overt subjects rose to 47% in *Eneas* (verse), but reached 81% in *Lapidaire* (prose), which is highly untypical rate for the time. Similarly, at the beginning of the thirteenth century, *Graal*, in prose, displayed 79% of overt subjects whereas this rate was only 68% in *RenartDole*, in verse. The discrepancy becomes even clearer when different passages in the same text are compared, as is possible in *Aucassin*, a text in which passages in verse alternate with passages in prose. The prose sections exhibit higher rates of overt subjects than the sections in verse, with 67% versus 48%, respectively.

The pioneering character of prose may well be a longstanding one. However this criterion is difficult to verify, as we have only a few prose texts before the thirteenth century, and most of them are in fact translations from Latin (in which null subjects were frequent), which may have influenced the style of the translator. This is the case of the *Quatre livres des Rois*, which is a late twelfth-century prose translation of the Bible.⁹ In this text, the frequency of overt subjects (58%) hardly exceeds that of contemporary texts in verse (*TroyesYvain* and *BeroulTristan*: respectively, 54% and 59% of overt subjects). The situation after the thirteenth century is quite different, and, so to speak, opposite, since verse becomes more specialized from Middle French on. From this point of view, the thirteenth century offers the best picture, as this period is balanced between verse and prose, at least for literary texts.

The domain (cf. section 4.2.3) also plays a major role but it largely overlaps the previous criterion, and it is difficult to disentangle the two criteria. While literary texts were written both in verse and in prose (at least from the thirteenth century on), non-literary texts were on the other hand almost all written in prose. Therefore we may wonder whether the modern and pioneering syntax of *Lapidaire*, in comparison with contemporary texts, results from its non-literary and didactic nature, or from the fact that it is written in prose, or from both, the two features (prose and non-literary/non-religious) still being infrequent at that time. The same question holds for *Beaumanoir* in the late thirteenth century.¹⁰ Similarly, the rate of overt subjects was higher in thirteenth- to fourteenth-century historical texts (*ClariConstantinople* and *Joinville*), but these texts were also written in prose. However, it is possible that in these texts the presence of multiple referents, which is inherent to historical writing, may have favoured the overt expression of subjects, for the sake of clarity. Likewise, the exceptionally high rates of overt subjects in legal charters probably result from the need to avoid any referential ambiguity. The domain and the form of texts seem to be still influential factors in the late fourteenth century. The frequency of overt subjects in *Griseldis*, a drama text in verse, rose to only 56%, versus 65% and 76% in *Mesnagier* and *Manieres*, two contemporary non-literary texts in prose.

It can be concluded from what precedes that both prose and the non-literary nature of the text favoured the propagation of overt subjects, though it is difficult to assess the exact contribution of each factor, as these two features are closely interrelated.

4.4.2 The dialectal parameter

⁹ Translations were excluded from the data in Table 4.1.

¹⁰ The relatively low general rate of overt subjects in *Mesnagier* (65%) results from the fact that this rate takes all types of clause into account, and this text includes a high number of imperative clauses, with null subjects.

The diatopic parameter does not seem to have been very influential in the expression of subjects. Admittedly, the data to support this statement is lacking: it is impossible to have a dialectal sampling before the thirteenth century, since Norman or Anglo-norman texts largely prevailed till the late twelfth century¹¹ Studies have been conducted on corpora of legal charters, which have the advantage of being reliably dated and located.¹² They have revealed only slight discrepancies, which, moreover, may differ from one study to another. Dees et al. (1980) thus pointed out that the rate of null subjects (at least in main declarative clauses) was higher in western dialects than in eastern ones. Schøsler (1984), on the other hand, showed that northern charters had the lowest rate of null subjects while Parisian ones had the highest.

4.4.3 Linguistic structures: types of clauses and argument structure

From the earliest texts on, the overt expression of subjects, and more specifically of personal pronouns, was far more frequent in subordinate clauses than in main declaratives (with no significant disparities between the different types of subordinates).¹³ Table 4.1 reveals a constant tendency during the whole period under scrutiny: the frequency of overt subjects in subordinates oscillated between 60% and 96% (except in *StLegier*: 48%) while in main declarative clauses it did not rise to 50% before the late twelfth century (with an exception in *Lapidaire*). The difference between the rates ranged from 20 to 30 percent if considering all the subjects, and exceeded 60 percent, in some texts, if considering only personal pronouns. The discrepancies persisted beyond the late fourteenth century, although, from then on, the difference tended to decrease and was no longer systematic, some texts even displaying reverse tendencies (*Manières*). Interestingly, the greatest discrepancies are observed from the mid-twelfth century till the late thirteenth century, i.e. during the first significant rise in overt expression.

Subordinate clauses thus appear to be pioneering in comparison with main declarative clauses. This observation has been interpreted in a number of ways. According to one family of proposals, the main-subordinate asymmetry is related to the fact that subordinate clauses generally resist verb-second (V2) syntax (for two analyses along those lines, couched in a generative framework, see Adams 1987; Vance 1997). However, such an explanation does not account for all types of subordinate clauses, and the main-subordinate asymmetry remains so far largely unexplained (see Prévost and Marchello-Nizia 2020: 1063-4). It should be mentioned that in Latin it was on the contrary subordinate clauses which more readily allowed for subject omission. The change most likely took place during Late Latin, but there is a serious lack of data to fill the gap and provide an explanation for this noticeable change.

The argument structure also seems to have played a role in the expansion of overt subjects. Studies (Rouquier and Marchello-Nizia 2012) on the argument structure in the oldest texts (*Passion*, *StLegier*, *StAlexis*, *Roland*) indeed showed that subjects were far less often expressed when a direct object was present, the difference ranging from 20 to 30 percent. In *Roland*, for instance, the frequency of overt subjects is 40% with transitive verbs, but it rises to 67% with intransitive verbs. In *TroyesYvain*, the respective frequencies are 37% and 66%. However, the difference decreased sharply from the beginning of the thirteenth century on, especially in prose (*Graal*: 77% versus 80%).

4.4.4 Direct speech versus narrative; verbal persons

¹¹ On the other hand, the dialectal factor is no longer relevant after the fifteenth century.

¹² See Balon and Larrivé (2016) for a study on the increase of overt subjects in legal texts.

¹³ This is a long-established observation: see among others Franzén (1939: 14-34); Adams (1987); Hirschbühler (1989; 1990); Prévost and Marchello-Nizia (2020: 1062-4).

Lastly, two parameters need to be considered, which, though pertaining to different levels of analysis – the situation of communication as an external factor, and the verbal person as an internal factor – deserve to be handled together since they are closely related.

If we distinguish between the first-person singular (P1) and third-persons singular and plural (P3), the most frequent ones, we observe quite significant differences in the frequencies of overt expression of the personal pronoun, at least in main declarative clauses, as is evidenced by the data in Table 4.2. Without denying the existing disparities between texts, it can be seen that, from the oldest texts, the frequency of overt expression of P3 is always lower than that of P1 (except in *StAlexis*). As early as the beginning of the thirteenth century, the frequency of overt P1 oscillated between 80% and 94% (except in *RenartDole*), while the overt expression of P3 was not systematically dominant before the early fourteenth century.

Table 4.2 Rates of expression of first- and third-person pronouns

Texts	Overt P1	Overt P3
<i>StAlexis</i> (c. 1050)	17%	24%
<i>Roland</i> (c. 1100)	34%	16%
<i>EneasI</i> (c. 1155)	24%	22%
<i>BeroulTristan</i> (late 12 th)	41%	25%
<i>TroyesYvain</i> (1177-1181)	51%	39%
<i>RenartDole</i> (early 13 th)	57%	52%
<i>ClariConstantinople</i> (after 1205)	82%	44%
<i>Graal</i> (c. 1225)	80%	60%
<i>Joinville</i> (1309)	80%	70%
<i>Manières</i> (1396, 1399)	94%	62%
<i>QuinzeJoies</i> (c. 1400)	83%	64%

These results imply that, from the early texts, in declarative clauses, ‘I’ (and in some texts ‘we’ and ‘you’) was expressed far more often than ‘he/she/they’. In contrast, we do not observe any significant difference between persons in the subordinate clauses of these same texts, which may be due to the fact that, even in early texts, the rate of overt expression in these clauses was very high. These data bear out the idea that a turning point occurred in the early thirteenth century, when the expression of first-person pronouns became prevalent in all types of clauses, and almost systematic in subordinate clauses, in which the prevailing expression henceforth concerned all persons.

But this skewing in the distribution of overt first- and third-person pronouns also results from the fact that P1 prevalently appears in direct speech. Focusing on the distribution of P3 pronouns in narrative and direct speech passages (in main declarative clauses), Table 4.3 shows that overt expression is always more frequent in direct speech:

Table 4.3 Rates of expression of first- and third-person pronouns in direct speech and in narrative¹⁴

	Overt P1	Overt P3	
	Direct speech	Direct speech	Narrative
<i>StAlexis</i> (c. 1050)	17% (15/89)	40% (10/25)	23% (72/311)
<i>Roland</i> (c. 1100)	34% (45/132)	24% (29/120)	11% (19/175)
<i>EneasI</i> (c. 1155)	24% (13/54)	39% (24/62)	21% (95/448)

¹⁴ I leave for a more extended study the introduction of chi squared tests.

<i>BeroulTristan</i> (late 12 th)	41% (47/115)	41% (41/100)	19% (48/255)
<i>TroyesYvain</i> (1177-81)	51% (39/77)	47% (41/88)	37% (135/362)
<i>ClariConstantinople</i> (after 1205)	82% (14/17)	60% (5/9)	44% (158/362)
<i>RenartDole</i> (early 13 th)	57% (54/95)	53% (46/86)	51% (149/290)
<i>Graal</i> (c. 1225)	83% (74/89)	81% (48/59)	56% (185/330)
<i>Manières</i> (1396, 1399)	94% (148/158)	66% (63/80)	58% (38/66)
<i>QuinzeJoies</i> (c. 1400)	82% (103/124)	81% (57/70)	60% (177/297)

However it is not so straightforward to disentangle the effects of different person contexts and those of the opposition between narrative passages and direct speech. There are probably two converging factors at work here. First, as they refer to the speaker, person pronouns were a privileged place for expressiveness and emphasis, which resulted in a higher rate of overt expression. Second, first-person pronouns prevalently occurred in direct speech, which may be considered as closer to the spoken language than narrative is. If we assume that spoken language may have been more advanced in its evolution, we may expect it to have displayed more overt pronouns, which may be reflected to a certain extent in direct speech.¹⁵ The increasing use probably resulted in rhetorical devaluation, which led to a widespread use – and thus weakening – of pronouns. The change clearly began with first-person pronouns, and then spread to the other persons (Detges 2003). The starting point was probably in (Late) Latin – though the precise date is debated.¹⁶

The preceding sections have highlighted that overt subjects initially increased more speedily in specific contexts: prose and non-literary texts, subordinate clauses, intransitive constructions, direct speech, and with first-person pronouns. These contexts thus appear to have been pioneering, and can be interpreted as factors of the propagation of overt subjects.¹⁷ We will see in the following section that these contexts played a far less influential role as regards the evolution of the position of the subject.

4.5 The increase in preverbal subjects

The evolution of the position of subjects offers a quite different picture, as regards both the general chronology and the factors which were influential in the progression of overt subjects. As for the expression of subjects, I focus here also on main declarative and subordinate clauses, since they are the most frequent ones, and also because the position of subjects was early highly constrained in the other types of clauses (interrogative, injunctive, exclamatory, and parenthetical clauses).

Although postverbal subjects, at least nominal ones, are not infrequent in Modern French, they have considerably receded since Old French, along paths which differ according to their nominal or pronominal nature,¹⁸ as is evidenced by Table 4.4.

¹⁵ Moreover, as the first-person pronoun already existed in Latin, it was, in some ways, a few centuries ahead of third-person pronouns.

¹⁶ See Pinkster (1987: 370); Touratier (1994: 23-5); Väänänen (1981: 123); Ernout and Thomas (1951: 143).

¹⁷ I leave aside here syntactic, semantic and pragmatic aspects, as it is not the purpose of this study (see Prévost and Marchello Nizia 2020: 1057-79 for a detailed account). Suffice it to say that null subjects first lost ground in non-coordinated structures, then in coordinated structures with different subjects and/or temporal orientation.

¹⁸ Among pronominal subjects only personal pronouns, the most frequent, are taken into account here. The others have always behaved more or less like nominal subjects, with the exception of indefinite *on*, which was assimilated to personal pronouns from the seventeenth century on.

Table 4.4 Rates of nominal and pronominal preverbal subjects in main declarative and subordinate clauses from the 11th to the 17th century¹⁹

Texts	Rate of preverbal nominal S in declaratives	Rate of preverbal pronominal S in declaratives	Rate of preverbal nominal S in subordinates	Rate of preverbal pronominal S in subordinates
<i>Passion</i> (c. 1000)	63% (73/116)	84% (31/37)	74% (20/27)	94% (17/18)
<i>StLegier</i> (c. 1000)	71% (30/42)	84% (16/19)	80% (4/5)	100% (20/20)
<i>StAlexis</i> (c. 1050)	38% (50/131)	78% (28/36)	67% (20/30)	100% (48/48)
<i>Roland</i> (1100)	53% (139/261)	81% (51/63)	53% (17/32)	100% (57/57)
Lapidaire (mid-12 th)	87% (104/120)	94% (130/138)	91% (31/34)	100% (77/77)
<i>Eneas I</i> (1155)	60% (105/176)	94% (44/47)	71% (39/55)	100% (85/85)
<i>BeroulTristan</i> (late 12 th)	70% (122/173)	84% (58/69)	71% (34/48)	99% (91/92)
<i>TroyesYvain</i> (1177-81)	59% (64/109)	72% (44/61)	75% (43/57)	98% (176/179)
<i>Aucassin</i> (late 12 th -13 th)	72% (84/116)	83% (117/141)	87% (48/55)	99% (129/130)
<i>ClariConstantinople</i> (after 1205)	43% (52/122)	57% (39/68)	98% (128/131)	97% (211/217)
<i>Graal</i> (c. 1225)	49% (51/105)	80% (90/113)	94% (66/70)	98% (259/263)
<i>RenartDole</i> (early 13 th)	57% (72/127)	78% (79/101)	87% (40/46)	99% (190/191)
<i>Beauma noir</i> (1283)	76% (63/83)	71% (53/75)	98% (171/174)	99% (255/258)
<i>Joinville</i> (1305-9)	67% (65/97)	92% (144/156)	98% (78/80)	100% (299/299)
<i>MachautFortune</i> (1341)	83% (60/72)	90% (103/114)	88% (80/91)	100% (160/160)
<i>Mesnagier</i> (1393)	83% (59/71)	97% (65/77)	97% (66/68)	100% (218/218)
<i>Griseldis</i> (1395)	56% (31/55)	89% (92/103)	71% (41/58)	100% (147/147)
<i>Manières</i> (1396, 1399)	66% (42/64)	97% (244/251)	88% (46/52)	99% (191/192)
<i>QuinzeJoies</i> (c. 1400)	63% (46/73)	94% (163/173)	99% (67/68)	99% (252/255)
<i>GersonSermon</i> (1402)	55% (48/87)	93% (115/123)	71% (63/89)	100% (142/142)
<i>Pathelin</i> (1456-69)	100% (42/42)	90% (277/307)	80% (12/15)	100% (163/163)
<i>CentNouvelles</i> (1456-67)	82% (75/90)	90% (120/133)	92% (89/98)	99% (238/239)
<i>ArchierBaignollet</i> (1468)	71% (17/24)	95% (112/118)	70% (7/10)	100% (42/42)
<i>Commynes</i> (1490-1505)	46% (75/162)	89% (72/81)	81% (82/101)	100% (183/183)
<i>JehanParis</i> (1494)	67% (70/104)	92% (114/124)	89% (88/99)	100% (182/182)
<i>Vigneulles</i> (1515)	73% (64/88)	99% (143/144)	91% (51/56)	99% (249/250)
<i>CalvinLettres</i> (1549)	100% (36/36)	97% (255/262)	90% (87/97)	99% (372/373)
<i>RonsardMisères</i> (1563)	91% (85/93)	100% (80/80)	94% (84/89)	100% (56/56)
<i>LéryBresil</i> (1578)	93% (53/57)	93% (168/180)	95% (106/112)	99% (296/297)
<i>Montaigne</i> (1592)	90% (104/117)	96% (186/194)	84% (84/100)	100% (196/196)
<i>BeroaldeParvenir</i> (1616)	96% (79/82)	98% (202/207)	84% (60/71)	99% (157/158)
<i>SorelBerger</i> (1627)	98% (91/93)	98% (188/192)	90% (96/107)	99% (312/315)
<i>Descartes</i> (1637)	100% (19/19)	100% (110/110)	90% (138/153)	100% (351/351)
Rate lower than 50%	Rate between 50 and 80 %		Rate higher than 80%	

Table 4.4 clearly shows that nominal and pronominal preverbal subjects did not expand at the same pace. As they already largely prevailed at the very beginning of French, preverbal

¹⁹ *Italics*: texts in verse; roman: texts in prose.

pronominal subjects did not undergo a very significant progression over the centuries. From more than 80% in most texts, they had reached more than 90% by the early fourteenth century, and their rate in subordinate clauses always oscillated between 94% and 100%, downgrading postposition to a very marginal phenomenon (1).

(1) Demi Espagne	vus durat	il	en fiet (<i>Chanson de Roland</i>)
O	V	S	adverbial
Half of Spain	you will give	he	as fief
‘He will give you half of Spain as yours’			

This situation can be accounted for by the informational principle which governed word order: as typical topics, pronominal subjects tended to appear in preverbal position. From the fourteenth century on, the contexts allowing for postverbal pronouns began to narrow, especially as regards initial elements, which were progressively restricted to adverbs with an epistemic value (*peut-être, à peine, ...*), as is still the case in Modern French (for a more detailed account, see Prévost and Marchello-Nizia 2020: 1102-13).

The very high rate of preverbal pronouns from the earliest texts onwards may explain why almost none of the factors that proved to be influential in the growing use of overt pronouns are relevant here. First, neither the form (verse or prose) nor the domain of the texts seems to have had any influence, as is evidenced by the comparison of two mid-twelfth-century texts: *Eneas* (verse and literary) and *Lapidaire* (prose and didactic) which display similar rates of preverbal pronominal subjects (94%). Likewise, in the early thirteenth century, the rates in *Graal* (prose) and *RenartDole* (verse) are very close. Second, we do not observe any significant difference as concerns direct speech/narrative opposition or the verbal person: depending on the texts, either first- or on the contrary third- person pronouns appeared more often in preverbal or in postverbal position. The only factor of variation is the type of clause, since subordinate clauses were slightly ahead of declarative clauses.

The postposition of nominal subjects has always been much more frequent:

(2) De Guanelun	atent	li reis	nuveles (<i>Chanson de Roland</i>)
Comp.	V	S	O
Of Guanelon	awaits	the king	news
‘The king awaits news of Guanelon’			

Moreover, postverbal subjects did not decrease in a linear fashion, as illustrated by Table 4.4. The rates of preverbal nominal subjects sometimes differed noticeably (at least in main declarative clauses) between contemporary texts, as is exemplified by *Aucassin* (72%) and *ClariConstantinople* (43%) in the early thirteenth century, or by *Commynes* (46%) and *JehanParis* (67%) in the late fifteenth century. It results from this that the rates were sometimes lower in some late fifteenth-century texts than in some late twelfth- or early thirteenth-century ones. The extension of preverbal nominal subjects during Medieval French was indeed very slow and sporadic, and it was only in the mid-sixteenth century that a turning point occurred: since then, preverbal subject rates have always been over 80%.

The decrease in postverbal nominal subjects was accompanied by progressive restrictions on verbs, in favour of intransitive ones, and on the elements likely to precede the verb. While they were nearly unconstrained until the thirteenth century, whether from a formal or functional point of view, the possible range then progressively decreased: arguments started to lose ground early on, and the range of adverbial elements narrowed sharply during the seventeenth century, to the advantage of spatial and temporal adverbs. The postposition of

nominal subjects is still well attested in Modern French, but it is restricted by the verb semantics and/or the weight of the subject and/or the presence of a locative element.

If we now turn to factors which proved to be influential in the growing use of overt subjects, we see that, as for pronominal subjects, only the type of clause seems to be relevant (albeit to a lesser extent than in the case of the expression of subjects). We observe in fact a constant difference in favour of subordinate clauses, which tended to increase from the early thirteenth century on, and declined sharply in the mid-sixteenth century, when both rates (declarative and subordinate clauses) of preverbal subjects reached over 90%. Apart from this factor, it is difficult to detect a clear tendency. There is no significant prevalence of preverbal subjects in direct speech, and the form of texts does not play a major part. For instance, in the early thirteenth century *ClariConstantinople* and *Graal*, both in prose, displayed a lower frequency of preverbal subjects than *RenartDole*, in verse. The domain has, for its part, a variable influence during the whole period: non-literary texts (*Lapidaire*, *Beaumanoir*, and *Mesnagier*) systematically displayed high rates of preverbal subjects, but literary texts behaved in a quite irregular way, with contrasted frequencies. A final point to mention is that there is no systematic correlation between the more or less advanced stage of a text as regards expression and its behaviour as regards the position of subjects: in comparison with contemporary texts, *Graal* displays a high rate of overt pronouns but a low rate of preverbal nominal subjects.

4.6 Conclusions and perspectives

The evolution of the syntax of subjects over the centuries has proceeded along differentiated chronologies and modalities, depending on whether one considers the expression or the position of subjects, and, in the latter case, the nature of the subjects.

We have seen that the two evolutions under scrutiny – increase in overt subjects and in preverbal subjects, either pronominal or nominal – display very different trajectories. Overt subjects, which were in a minority in declarative clauses at the beginning of French, went through a first rise in the early thirteenth century, then through a second one in the mid-sixteenth century. As for the position of subjects, pronominal ones had always prevailed in preverbal position, and therefore increased only slightly. The evolution of nominal subjects is the most complex one, as it was very irregular until the mid-sixteenth century, the Modern French situation being reached only in the seventeenth century.

Considering the question of expression as a starting point, I have identified a number of internal and external contexts which have proved to be factors conducive to the progression of overt subjects. On the other hand, most of these factors proved to have little – if any – influence as regards the evolution of the position of subjects.

The only factor common to all these changes is the pioneering nature of subordinate clauses, albeit with some nuances: there is a large discrepancy between the rates in declarative and subordinate clauses as regards overt subjects, while it is lower for preverbal pronominal subjects (due to the fact that, very early on, their frequency was very high in declarative clauses) and variable for preverbal nominal subjects. I do not have a truly satisfactory explanation for this observation, which moreover contradicts the idea that subordinate clauses are generally conservative (Vennemann 1975; Givón 1979), and which is all the more surprising since, in Latin, overt subjects were on the contrary less frequent in subordinate clauses. This question needs to be further investigated.

As to the other factors, two hypotheses may be formulated to account for their variable influence. The first is a quantitative one: it seems that, in general, the factors were no longer very influential once the change was advanced, that is as soon as one of the variants had – consistently through the texts – overtaken the other one. This may explain why not all the factors had an effect, on the one hand, in subordinate clauses (as overt and preverbal subjects had always largely prevailed in them), and, on the other hand, on the increase in preverbal

pronominal subjects, which prevailed in all contexts from the very beginning of French. This hypothesis is corroborated by the evolution of the position of nominal objects, which took place very early on (see Marchello-Nizia 1995): the reversal in frequencies of OV and VO occurred between the late ninth century and the early thirteenth century, which makes it difficult to analyze the change in detail since we can only rely on sparse data for the decisive period (and it is most likely that the change had already begun in Late Latin). If we look at the data in the twelfth to thirteenth centuries, it appears that only the opposition between direct speech and narrative was influential. VO order was more frequent in direct speech, although the difference was very slight (1-10%), and frequencies tended to level out from the end of twelfth century on.²⁰ It is not very likely that this factor played a major role, at least at that time. As regards the opposition between main declarative and subordinate clauses, no clear tendency can be discerned. VO order (versus OV) was more frequent either in declarative or in subordinate clauses, depending on the texts, even in the most remote period. This factor, however, may have been influential earlier, in Late Latin, at the beginning of the change.

The question arises on the contrary as to the increase in preverbal nominal subjects. It is clear that prose did not offer a pioneering context, as it did for overt subjects. As for the domain, things are less straightforward: until the sixteenth century non-literary texts (in particular didactic ones) displayed fairly high rates of preverbal subjects, but no clear tendency can be observed in literary texts. On the other hand, in Modern French, descriptive passages, whatever the textual domain, tend to favour postverbal subjects,²¹ their usage pertaining most often to stylistic effects.²²

The fact that the direct speech/narrative opposition strongly influenced the increase in overt subjects but had no effect on the increase in preverbal subjects may be accounted for by the fact that overt expression (especially first-person pronouns) was associated with expressivity and emphasis, which are more likely to appear in spoken language, and hence in direct speech (as it is a kind of reflection of it). No expressive effect has ever been associated with the preverbal position, and it is thus no surprise that direct speech was not a pioneering context for the propagation of preverbal contexts.

I am well aware that the preceding pages have sketched an incomplete picture of the role that different factors played in the increase in overt subjects on the one hand, and of preverbal subjects on the other. There still remain many aspects to investigate to provide a more thorough account. However I hope to have shown that taking these different factors into account makes it possible both to refine the chronology of these changes and to (partly) suggest some motivations for the changes, while identifying the contexts (in the broad sense) conducive to these evolutions.

Be it recalled that this contribution does not claim in any way to provide a true socio-historical study, but, more modestly, some kind of a variationist approach, which may, hopefully, bring new insights as regards two major and related changes which took place in the history of French.

Appendix 1: References of the texts of the corpus

[ArchierBaignollet] *Franc (Le) archier de Baignollet* (Fr)

[Aucassin] *Aucassin et Nicolette* (BFM)

[Beaumanoir] Philippe de Beaumanoir, *Coutumes de Beauvaisis* (BFM)

[BeroaldeParvenir] Béroalde de Verville, François, *Le Moyen de parvenir* (Fr)

[BeroulTristan] Beroul, *Tristan* (BFM)

[CalvinLettres] Calvin, Jean, *Lettres à Monsieur et Madame de Falais* (Fr)

²⁰ Personal unpublished research.

²¹ Though only in a relatively narrow set of contexts and the phenomenon quantitatively remains in the minority.

²² See Carruthers (2018).

[CentNouvelles] *Cent nouvelles nouvelles* (Fr)
 [ClariConstantinople] Robert de Clari, *Conquête de Constantinople* (BFM)
 [Commynes] Philippe de Commines, *Mémoires* (BFM)
 [Descartes] Descartes, René, *Discours de la méthode* (Fr)
 [Eneas] *Eneas 1 et 2* (BFM)
 [GersonSermon] Gerson, Jean, *Sermon pour la fête de la Sainte Trinité*. (Fr).
 [Graal] *Queste del saint Graal* (BFM)
 [Griseldis] *Estoire de Griseldis en rimes et par personnages* (BFM)
 [JehanParis] *Roman de Jehan de Paris* (BFM)
 [Joinville] Jean de Joinville, *Vie de saint Louis* (Fr)
 [Lapidaire] *Lapidaire en prose* (BFM)
 [LéryBrésil] Léry, Jean de, *Histoire d'un voyage fait en la terre du Bresil*. (BVH)
 [MachautFortune] Guillaume de Machaut, *Remede de Fortune* (BFM)
 [Manières] *Manières de langage* (BFM)
 [Mesnagier] *Mesnagier de Paris* (BFM)
 [Montaigne] Montaigne, Michel de, *Essais* (Fr)
 [Passion] *Passion de Jésus-Christ ou Passion de Clermont* (BFM)
 [Pathelin] *La Farce de maître Pierre Pathelin* (BFM)
 [QuinzeJoies] *Quinze joies de mariage* (BFM)
 [RenartDole] Jean Renart, *Roman de la Rose ou de Guillaume de Dole* (BFM).
 [Roland] *Chanson de Roland* (BFM)
 [RonsardMisères] Ronsard, Pierre de, *Discours des Misères de ce temps* (BVH)
 [SorelBerger] Sorel Charles, *Le Berger extravagant* (Fr)
 [StAlexis] *Vie de saint Alexis* (BFM)
 [StLegier] *Vie de saint Léger* (BFM)
 [TroyesYvain] Chrétien de Troyes, *Chevalier au Lion ou Yvain* (BFM).
 [Vigneulles] Vigneulles, Philippe de, *Les Cent Nouvelles nouvelles* (Fr)

Linguistic references

Extended references can be found in the databases the texts stem from:

BFM – *Base de Français Médiéval* (2012-), Lyon, UMR ICAR (CNRS & ENS de Lyon), <http://bfm.ens-lyon.fr>. (BFM). Accessed November 2020.

Bibliothèques virtuelles humanistes <http://www.bvh.univ-tours.fr>, 2009 (BVH).

FRANTEXT – *Base textuelle Frantext*, ATILF (CNRS & Université de Lorraine), <http://www.frantext.fr>. (Fr)

Adams, Marianne (1987). *Old French, Null Subjects and Verb Second Phenomena*. Ph.D. Dissertation. University of California, Los Angeles.

Andersen, Henning (2001). 'Actualization and the (uni)directionality of change'. In H. Andersen (éd.) *Actualization. Linguistic Change in Progress*. Amsterdam: John Benjamins Publishing Company, 225-48.

Ariel, Mira (1988). 'Referring and accessibility', *Journal of Linguistics* 24, 65-87.

Ayres-Bennett, Wendy (1994). 'Negative evidence, or, Another look at the non-use of negative *ne* in seventeenth-century French', *French Studies* 48, 63-85.

Ayres-Bennett, Wendy (1996). *A History of the French Language through Texts*. London: Routledge.

- Ayres-Bennett, Wendy (2004). *Sociolinguistic Variation in Seventeenth-Century France. Methodology and Case Studies*. Cambridge: Cambridge University Press.
- Balon, Laurent et Pierre Larrivée (2016) 'L'ancien français n'est déjà plus une langue à sujet nul—nouveau témoignage des textes légaux', *Journal of French Language Studies*, 26(2): 221-237
- Carruthers, Janice (2018). 'Postposition of the Subject in Contemporary French. An Exploration of Medium, Register and Genre', *Modern Languages Open* 2018 (1).12, 1-20, online, DOI: <http://doi.org/10.3828/mlo.v0i0.189>
- Combettes, Bernard (1988). *Recherche sur l'ordre des éléments de la phrase en moyen français*. Thèse pour le Doctorat d'Etat. Université de Nancy.
- Dees, Anthonij *et al.* (1980). *Atlas des formes et des constructions des chartes françaises du 13e siècle*. Tübingen : Max Niemeyer Verlag (Beihefte zur Zeitschrift für romanische Philologie. 178).
- Detges, Ulrich (2003). 'Du sujet parlant au sujet grammatical. L'obligatorisation des pronoms sujets en ancien français dans une perspective pragmatique et comparative', *Verbum* 25 (3), 307-33.
- Ernout, Alfred and Thomas, François (1951). *Syntaxe latine*. Paris: Klincksieck.
- Firbas, Jan (1992). *Functional Sentence Perspective in Written and Spoken Communication*. Cambridge: Cambridge University Press.
- Foulet, Lucien (1930 [1919]). *Petite syntaxe de l'ancien français*. Paris: Champion.
- Foulet, Lucien (1935). 'L'extension de la forme oblique du pronom personnel en ancien français'. *Romania* 61 (243 and 244), 257-315 and 401-63.
- Franzén, Torsten (1939.) *Étude sur la syntaxe des pronoms personnels sujets en ancien français*. Upsal: Almqvist et Wiksells.
- Gadet, Françoise (1997) 'La variation, plus qu'une écume', *Langue Française* 115: 5–18.
- Givón, T. (1979). *On understanding grammar*. New York: Academic Press.
- Hirschbühler, Paul (1989). 'On the Existence of Null Subjects in Embedded Clauses in Old and Middle French', in C. Kirschner and J. de Cesaris (ed.), *Studies in Romance Linguistics*. Amsterdam: John Benjamins, 155-76.
- Hirschbühler, Paul (1990). 'La légitimation de la construction V1 en subordonnée dans la prose et le vers en ancien français'. *Revue Québécoise de Linguistique* 19 (1), 33-55.
- Jakobson, Roman (1963). *Essais de linguistique générale*, vol. 1. Paris: Éditions de Minuit.
- Labov, William (1972). 'Some Principles of Linguistic Methodology'. *Language in Society* 1 (1), 97-120.
- Lambrecht, Knud (1994). *Information Structure and Sentence Form. Topic, Focus and the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press.
- Lodge, R. Anthony (1993). *French: From Dialect to Standard*. London: Routledge.
- Lodge, R. Anthony (2004). *A Sociolinguistic History of Parisian French*. Cambridge: Cambridge University Press.
- Marchello-Nizia, Christiane (1995). *L'évolution du français. Ordre des mots, démonstratifs, accent tonique*. Paris: Armand Colin.
- Marchello-Nizia, Christiane (2020). 'L'objet', in C. Marchello-Nizia, B. Combettes, S. Prévost and T. Scheer (eds) *Grande Grammaire Historique du Français*, Berlin: De Gruyter Mouton, 1126-56.
- Meillet, Antoine (1916). Compte-rendu du Cours de linguistique générale de F. de Saussure. *Bulletin de la société linguistique de Paris*, 64, 32-36.
- Pinkster, Harm (1987). 'The strategy and chronology of the development of Future and Perfect Tense auxiliaries in Latin'. In M. Harris and P. Ramat (eds) *Historical Development of Auxiliaries*. Berlin: Mouton de Gruyter, 193-223.
- Posner, Rebecca (1997). *Linguistic Change in French*. Oxford: Oxford University Press.

- Prévost, Sophie (2018). 'Increase of pronominal subjects in Old French: Evidence for a starting-point in Late Latin', in A. Carlier and C. Guillot (eds), *Latin tardif - français ancien: continuités et ruptures*, coll. Beihefte zur Zeitschrift für romanische Philologie. Berlin: De Gruyter, 171-200.
- Prévost, Sophie, and Monique Dufresne (2020). 'L'approche du changement linguistique dans la GGHF', in C. Marchello-Nizia, B. Combettes, S. Prévost and T. Scheer (eds) *Grande Grammaire Historique du Français*, Berlin: De Gruyter Mouton, 15-36.
- Prévost, Sophie, and Christiane Marchello-Nizia (2020). 'Le sujet', in C. Marchello-Nizia, B. Combettes, S. Prévost and T. Scheer (eds) *Grande Grammaire Historique du Français*, Berlin: De Gruyter Mouton, 1055-126.
- Roberts, Ian (2014). 'Taraldsen's Generalisation and Diachronic Syntax: Two Ways to Lose Null Subjects'. In P. Svenonius (ed.) *Functional Structure from Top to Toe: The Cartography of Syntactic Structures*. Oxford: Oxford University Press, 115-48.
- Romaine, Suzanne (1982). *Socio-historical Linguistics: Its Status and Methodology*, Cambridge: Cambridge University Press.
- Romaine, Suzanne (1988). 'Historical sociolinguistics: problems and methodology', in U. Ammon, N. Dittmar and K. J. Mattheier (eds), *Sociolinguistics: An International Handbook of the Sciences of Language and Society*, vol. 2. Berlin / New York: W. de Gruyter, 1452-69.
- Rouquier, Magali, and Christiane Marchello-Nizia (2012). 'De (S)OV à SVO en français: où et quand? L'ordre des constituants propositionnels dans la *Passion de Clermont* et la *Vie de saint Alexis*'. In M. Dufresne (ed.) *Constructions en changement. Hommage à Paul Hirschbühler*. Laval: Presses de l'Université de Laval, 111-55.
- Schøsler, Lene (1984). *La déclinaison bicasuelle de l'ancien français, son rôle dans la syntaxe de la phrase, les causes de sa disparition*. Odense: Odense University Press (Etudes Romanes de l'Université d'Odense, 19).
- Schøsler, Lene (2002). 'La variation linguistique: le cas de l'expression du sujet'. In R. Sampson et W. Ayres-Bennett (ed.) *Interpreting the History of French. A Festschrift for Peter Rickard on the occasion of his eightieth birthday*. Amsterdam / New York: Rodopi, 187-208.
- Simonenko, Alexandra, Benoît Crabbé, and Sophie Prévost (2019). 'Agreement syncretisation and the loss of null subjects: quantificational models for Medieval French', *Language Variation and Change* 31 (3), 275-301.
- Skårup, Povl (1975). *Les premières zones de la proposition en ancien français. Essai de syntaxe de position*. Etudes romanes de l'université de Copenhague, *Revue Romane numéro spécial* 6. Copenhague: Akademisk Forlag.
- Touratier, Christian (1994). *Syntaxe Latine*. Louvain-la-Neuve: Peeters.
- Väänänen, Veikko (1981). *Introduction au latin vulgaire*. Paris: Klincksieck.
- Vance, Barbara (1997). *Syntactic Change in Medieval French. Verb-Second and Null Subjects*. Dordrecht / Boston / Londres: Kluwer Academic Publishers.
- Vennemann, Théo (1975). 'An explanation of drift', in C. Li (ed.) *Word order and word order change*. Austin: University of Texas Press, 267-305.
- Vennemann, Theo (1976). 'Topics, subjects and word-order: from SXV to SVX via TVX', in J. M. Anderson and C. Jones (eds) *Proceedings of the first International Congress of Historical Linguistics*. Amsterdam: John Benjamins, 339-76.