

Syntactic Glosses and Historical Syntax*

Michelle Troberg & John Whitman

1 Introduction

This paper explores the significance of syntactic glosses as a data source for historical syntax, and the specific utility of a glossed 12th century text for the study of Old Occitan (OOc). We show that in addition to their importance as sources of information about phonology and lexicon, glossed materials are a potential source of information about the syntax of languages in periods when extensive vernacular materials are scarce or lacking. We also make the point that analysis of syntactic glosses requires a theoretically informed understanding of the grammar under investigation, a point first established in the generative study of older Indo-European languages by the scholar whom we honor (Hale 1987).

The target text is MS Paris, Bibliothèque Nationale, fonds latin 7979 (BnF 7979), previously noted by Munk Olsen (1982, 1996, 2007, 2009) and studied by Reynolds (1990a, 1990b, 1996, 2000). BnF 7979 contains the *Odes, Ars poetica et Satires* of Horace, as well as Lucan's *Pharsalia*. Its precise provenance is uncertain, but all authors agree that it was produced in the south of France in the 12th century. The ms. is extensively glossed with sequence glosses, a specific type of syntactic gloss or construe mark using letters of the alphabet, written interlinearly above words or phrases in the text. The lemmata could then be re-arranged in alphabetic order, for a purpose whose precise identity is a major focus of the paper.

2 Syntactic glossing

BnF 7979 is remarkable for the extensiveness of its syntactic glossing: Essentially every line has sequence glosses, from the first *Ode* through the *Pharsalia*. This makes it a text well-suited to cast light on a longstanding dispute in research on medieval glossed texts: the possible connections between syntactic glossing and contemporary vernaculars. Draak (1967) first proposed that the orders imposed by syntactic glosses on Latin manuscripts glossed by Irish monks indicated that the texts were to be read, or if one prefers, translated, according to Irish syntax. Draak pointed to the heavy preponderance of verb-initial word order in the glosses she studied. A similar

* We thank two reviewers, Bernhard Koller, and George Walkden, as well as Franck Cinato and Christine Meklenborg for their valuable comments on earlier drafts. All errors are of course our own.

Whitman's research on this paper was supported by the Laboratory Program for Korean Studies through the Ministry of Education of the Republic of Korea and the Korean Studies Promotion Service of the Academy of Korean Studies (AKS-2016-LAB-2250004).

argument is made by Robinson (1973) for sequence glosses on Latin texts of Anglo-Saxon provenance such as the Lambeth Psalter. Robinson’s argument is further supported by O’Neill (1992). Lemoine (1994) makes an argument similar to Draak’s for texts of Breton provenance. We will call this view, that the sequence glosses in a text in some way reference the vernacular language of the glossator, the vernacular reading hypothesis. Among glossed medieval texts associated with Romancespeaking regions, the so-called *Glosas Emilianenses*, glosses inserted in 11th century Spain in a collection of Latin sermons, have also been associated with vernacular reading (Wright 1982).

An opposed view, championed by Korhammer (1980), holds that the word order generated by sequence glosses is instead a variety of Latin known as the *ordo naturalis*, which corresponds to medieval grammarians’ idealized conception of how Latin should be put together, or understood. Reynolds (1990a), the first study of the BnF 7979 glosses, observed that “some features of the converted word order suggest a strong Occitan influence” (1990a: 37), but also suggests the word order generated by the glosses in certain respects “runs counter to it” (39).

We show that the word order of the glosses is fully consonant with the picture that has emerged in recent research of Old Occitan as a “relaxed” V2 language (Vance et al. 2009, Donaldson 2016, Wolfe 2018, 2019). We argue that this in turn supports the vernacular reading hypothesis for this particular text, and that, further, BnF 7979 and glossed Latin texts like it can be used to expand the corpus of early Romance vernacular materials, much as glossed Chinese texts in East Asia have been used to expand the corpus of earlier Japanese and Korean.

2.1 Syntactic glossing in East Asia

In East Asia, sophisticated systems of glossing were developed to assist in the reading of Literary Sinitic texts in the local vernacular (see Whitman et al. 2010 for Japanese and Korean and Kosukegawa and Whitman 2018 for Vietnamese). In Japan, research on this glossing system, known as *kunten* 訓点 ‘exegesis marks’ has been a major focus of historical linguistics for the past century. In Korea, research on *kugyŏl*, 口訣, ‘oral mnemonics’, the system that was probably the source for its Japanese counterpart, has blossomed since the 1970s.

Just as BnF 7979 glosses a canonical Latin text, Japanese *kunten* and Korea *kugyŏl* gloss texts in Classical Chinese (Literary Sinitic). Just as indisputably, the glossing generates a vernacular reading, a rendition of the Chinese lemmata in Japanese or Korean. Because such texts are available for periods when vernacular writing is sparse or difficult to interpret, they have emerged as important sources of information about the earlier syntax and morphology of both languages. A good example of this comes from recent research on Korean *kugyŏl* texts. This research has shown that the specialized adnominal endings for verbs and adjectives in Middle and Modern Korean originally functioned as VP and clausal nominalizers (Nam 2010), much like so called participial endings in other languages in the region. This fact was hinted at by various fossilized usages in Middle Korean, but could not be textually confirmed until *kugyŏl* evidence became available.

In this case, research on glossed texts opened a window on a period where other kinds of textual evidence are scant or unavailable: Korean prior to the invention of

hangul writing in the 15th century. The period we examine here, Old Occitan in the 12th century, has a somewhat similar status: Although verse texts exist prior to this time, prose texts are limited.

2.2 Syntactic glossing in the medieval West

Wieland (1983) proposes the following typology of glosses in medieval manuscripts:

Table 1: Wieland’s (1983) classification of glosses by function

(a)	Prosodic glosses	Metrical information: accent or syllable length.
(b)	Lexical glosses	Lexical equivalents for a word or phrase.
(c)	Grammatical glosses	Morphosyntactic information: case, part of speech, etc.
(d)	Syntactical glosses	Syntactic information (mainly word order).
(e)	Commentary glosses	Summarize content, give etymologies, etc.

“Syntactic(al) glosses,” also known as construal or construe marks, fall into two categories. The first are sequence glosses of the sort provided in every line of BnF 7979. The second are symbol glosses, which typically relate two or more noncontiguous terms in the Latin text. Examples of both of these are found in texts such as the 9th century Saint Gall manuscript of Priscian’s Latin grammar studied by Hoffman and accessible in the online resource developed by Bauer et al. (2017). The syntactic glosses in BnF 7979, in contrast, are only letter glosses.

2.3 The text: Horace’s Satires, BnF 7979

Both the Horace and Lucan portions of BnF 7979 have alphabetic sequence glosses on nearly every line. The manuscript also contains lexical and grammatical glosses, in Latin and Old Occitan. The *Odes* portion of the ms. is accompanied by neumatic or musical notation (Draheim and Wille 1985, Riou 1990, 1991). As Riou (1990: 265) points out, the interlinear sequence glosses force the scribe to enter the musical notation at the bottom of the page, where the text is reinscribed. The presence of the neumatic notation indicates that the scribe was aware not just of the textual content of Horace’s verse, but of its metrical and musical nature. Reynolds (1990a: 35) estimates that the syntactic glossing is “contemporaneous with the text,” and this is our judgment about the neumatic notation too.

We can get a sense of how all this works by looking at the first line of the *Carmina* (*Odes*) in the manuscript, as represented in Table 2.

Table 2: Poem 1, line 1 of the *Carmina* (*Odes*) in BnF fonds latin 7979, f.1r–1, l.1.

Glosses	a	<i>nate</i>	c	b	d
	<i>hi</i>	<i>o</i>	<i>natis</i>	<i>avis</i>	<i>nat</i>
Lemma	Mecenas		attavis	edite	regibus
	Maecenas.VOC		ancestor.ABL.PL	issued.from.M.SG.VOC	king.ABL.PL
	“O Maecenas, born of royal ancestors”				

The bold-faced lower-case letters are the sequence glosses. *M<a>ecenas*, and *edite*, the participle modifying it, are in the vocative, which is regularly glossed with *o*, a supplement gloss in the classificatory scheme of Blom (2017: 32–33); here, we regard it as part of the vernacular reading. The participial phrase, consisting of elements b, c, d, is reordered so that the participle *edite* immediately follows *Mecenas*. *Nat* is a Latin lexical (synonym) gloss on *edite*, abbreviating *natus* ‘born’. *Attavis* is accompanied by the lexical gloss *avis*, showing that the word is a compound involving *avis* ‘grandfather, ancestor’. The masculine demonstrative pronoun *hi* < *hic* ‘this’ indicates the gender of *M<a>ecenas*; Reynolds (1990a: 35, n.5) notes that this strategy for indicating gender is attested elsewhere in materials from Southwestern France. The declension endings *-nate* (ablative singular) and *-natis* (genitive singular), show that *M<a>ecenas* is a third declension noun. These first four lines are then rewritten at the bottom of the page without glosses but with neumatic notation added. The notation, consisting of a series of dots written above and below the line, indicates breath groups and pitch. In our transcription of the manuscript lines, we include the full form of the Latin words when, in the manuscript, they are abbreviated. We indicate any letters omitted in the manuscript with italics.

Throughout BnF 7979, lexical glosses in the vernacular as well as Latin are inserted to signal the meaning of less familiar Latin words and in some cases to indicate discourse referents. For example, the participle *demissis* ‘dropped, put down’, (Book 1, Satire 2, l. 25; BnF 7979 f. 17v, l.14; see example (10) below) is glossed with the OOc particle *abas* ‘down’.

3 Old Occitan syntax

3.1 Major constituent order in main clauses

Recent research argues that Old Occitan had a verb-second (V2) grammar (Benincà and Poletto 2004; Vance et al. 2009; Donaldson 2015, 2016; Wolfe 2016, 2018, 2019; but see Sitaridou 2012 for a different view), along with other early Romance languages. In matrix clauses, this means that the dominant word order places the finite verb in second position. However, compared to modern Germanic V2 languages such as German, and Old French among the better studied early Romance languages, OOc was a “relaxed” V2 language, in that it allowed verb-initial (V1), V3, V4, and very occasionally V5 and V6 orders as well.

Within a generative framework, these facts can be described as a grammatical requirement for the finite verb in Old Occitan to raise to a position in the articulated left periphery, the layer of syntactic representation where information structure is primarily encoded. Once the verb raises, another constituent, traditionally and pretheoretically called the “topic” (regardless of its actual information structural role) merges to its immediate left. This constituent needn’t be the subject: fronted objects and adjuncts also satisfy V2. The result, when a nonsubject topic is fronted to its immediate left, is that the verb in second position precedes the subject. This XP-V-NP_{subject} pattern, traditionally called “inversion”, is the hallmark of V2 languages and often claimed to be a crucial cue in the first language acquisition of V2 syntax. V2 with nonsubject topics is a distinguishing feature of Old Occitan syntax and, as we show below, the word order generated by the BnF 7979 glosses.

3.2 Major constituent order in subordinate clauses

Wolfe (2018) reports significant asymmetries between word order patterns in matrix clauses and those in embedded clauses. While only 23.78% of matrix clauses are subject initial, subject-initial clauses make up 76.47% of embedded contexts, as in (4):

- (4) *E qan Betrans de Born saup **qe-l** reis devia essir de preison*
 and when Bertran de Born knew that=the king should leave.INF of prison
 “and when Bertran de Born found out that the king was to leave prison”
 (*Vidas*, 80,8; Wolfe 2018: 9)

Wolfe points out that internal arguments can be fronted in embedded clauses, but, in his corpus, they only occur when the subordinate is the complement of a bridge verb (*say, respond, hear, know, show, seem*), which plausibly selects complements with matrix-like left peripheries. The evidence thus suggests that embedded clauses are generally small in OOC; in the majority of cases, the complementizer occupies the head of CP, preventing the verb from raising beyond TP.

3.3 Personal pronouns

OOC is a null subject language.¹ Overt subject pronouns, when they occur, are commonly used to mark focus (5a, where *nos* ‘we’ is strengthened by *autres*) or contrastive topic (5b), and can occur in coordination (5c).

- (5) a. *que **nos autres** desse venem*
 that we others soon come.1PL.PST
 “for *we* have just arrived” (Appel 9, 66; Jensen 1986: 87)
- b. ***nos** fom auster et **ylh** foro aigro*
 we were.1PL hawks and they were.3PL herons
 “*we* were hawks and *they* were herons”
 (R. de Vaqueiras: *Epic Letter II* 54; Jensen 1986: 87)
- c. ***vos** et **eu** avem lo pretz de tota l’altra gen*
 you and I have.1PL.PRES the praise of all the-other people
 “you and I are praised above everybody else”
 (P. Vidal XXIV, 60; Jensen 1986: 87)

Object pronouns are of two types: weak, clitic forms that occur directly before (6a) or after (6b) the verb, and strong forms. Strong forms are required in stressed positions, such as the complement of a preposition, contrastive contexts (6c), and frequently before non-finite verbs. Jensen (1986: ch. 5) notes, however, following Sneyders de Vogel (1927) that in many contexts the use of strong forms is not syntactically constrained in the strict sense; rather, subtle discourse and stylistic considerations can license a strong form when one might normally expect a clitic.

¹ As opposed to Old French, for example, which is a partial pro-drop language; see Sitaridou 2004.

- (6) a. *per eveja lo=mesdren e preiso*
 by envy him=put.PST.3PL in prison
 “out of envy they put him in prison”
 (Boeci: Appel 105,27; Jensen 1986: 95)
- b. *Et illi servi=lo am gran Misericordia.*
 and she served.3SG-him in great pity
 “And she served him with great pity” (Douceline, 46; Wolfe 2018: 5)
- c. *ie-ls pretz aitan pauc quon ilh me*
 I-them value as.much little than they me
 “I appreciate them as little as they do me”
 (Sordel XXVI 42; Jensen 1986: 101)

The conditions governing proclitics (6a) versus enclitics (6b) have been long debated. Donaldson (2016) argues that enclitics are possible only when the specifier position of the projection hosting the verb (for him, FocP) is empty. This has the consequence that any subject, including a subject pronoun, preceding the verb followed by an enclitic pronoun must be left dislocated to a position higher than FocP. Looking at our enclitic example, (6b), this is consistent with the fact that the subject in (6b) is a strong pronoun, a left dislocatable item.

3.4 OV Word Order

A striking difference between the main clause word order of Latin and that of modern Romance varieties is the scarcity of SOV order in the latter. Early Romance texts suggest that the change from OV to VO word order was still underway in the 12th and 13th centuries. Marchello-Nizia (1995) and Zaring (2010) point out that in Old French, SOV involving a finite verb is rarely attested in 12th c. verse (between 2.4% and 4.6%), and is all but absent in 13th century prose. However, the near-absence of OV in finite clauses can be accounted for by V raising, to C in matrix and to T in embedded clauses. Zaring presents data that show substantial OV word order through the 13th century involving nonfinite forms, such as past participles (43% in the 12th c. and 28% in the 13th c.) and infinitives (60% in the 12th c. and 47% in the 13th). These facts lead Zaring (2011) to propose that Old French is typologically OV. Zaring adopts the antisymmetric approach to OV order proposed by Biberauer and Roberts (2005) and developed specifically for Yiddish by Wallenberg (2009), deriving head-final surface order through a combination of head and phrasal movement with pied-piping. Zaring suggests that the already robust evidence for co-existing VO orders in the 12th century Old French data is the reflex of a competing grammar lacking the features driving movement to the left of the head. An alternative that comes to mind as part of a broader program to explain the change from OVAux Latin to Aux VO Romance is these features, particularly the EPP features driving phrasal movement and pied piping, first became optional and then were eventually lost in the daughter Romance languages.

Although a full investigation of OV order in Old Occitan is beyond the scope of this paper, the secondary literature suggests that OOc may have had a propensity for OV order in nonfinite clause similar to Old French. Jensen (1986) reports that while SOV order involving a finite verb, as in (7a), is rare, “(t)he two elements of a

periphrastic verb form may be separated”, by a variety of elements including object noun phrases, as in the participial example (7b) and the infinitive example (7c).

- (7) a. *o Deus tan rica saint' aduz*
 where God so rich saint bring.3SG.PRES
 “where God brings such a powerful saint”
 (*Sainte Foi*, v.438; Jensen 1986: 387)
- b. *el a son escut demandat*
 he AUX his sword ask.PST.PTCP
 “He asked for his sword” (*Jaufré*, v.208; Jensen 1986: 396)
- c. *comensei a chanso far desse*
 began to song make.INF immediately
 “I began immediately to write a *cansó*”
 (*P. Vidal XII*, 7; Jensen 1986: 98)

4 Method

We chose to analyse the first 54 lines of *Satire I-2* in BnF 7979, as this *Satire* contains a variety of clause types: declaratives, interrogatives, reported speech, quotative speech, etc. We also analysed 26 lines of *Satire I-1*, based on several lines that Reynolds (1990b) explicitly discusses and on lines that contained occurrences of object pronouns.² Finally, we analysed 59 lines from *Satires 2* through 10 in which we identified object pronouns.³ In total, we therefore examined 113 lines of Horace with regards to the syntactic glosses.

The transcription of the lines and their syntactic glosses was based entirely on the electronic version of BnF 7979, available through the BnF Richelieu (copied from the microfilm version). The quality of the electronic version varies, but in general, it offers a clear enough image to identify most of the syntactic glosses and punctuation and to a lesser extent the prosodic, lexical, grammatical, and commentary glosses. We used Fairclough’s translation of the *Satires, Epistles and Ars Poetica*, published in The Loeb Classical Library collection (Fairclough 1942), as a guide for interpreting the text.

5 Results and Discussion

5.1 Reynolds 1990a, 1990b

Before discussing our own findings, we first review those reported in Reynolds (1990a, 1990b, 1996), the only other examination of syntactic glosses in BnF 7979. Reynolds concludes that the glossed word order in BnF 7979 is, on almost all counts, consistent

² We can only be sure that Reynolds analysed *Satire I-1*, 1.54–56, as she discusses these lines in her 1990b paper. We found object pronouns in the following lines: **Satire I-1**, lines 1, 2, 3, 7, 8, 9, 10, 11, 12, 14, 19, 38, 39, 40, 41, 54, 55, 56, 66, 78, 81, 82, 83, 84, 120, 121.

³ **Satire I-2** lines 1–54, 57, 59, 60, 69, 70, 71, 77, 96, 97, 100, 103; **Satire I-3** lines 21, 22, 33, 34, 63, 123, 124; **Satire I-4** lines 14, 17, 18, 39, 70, 90, 93, 94, 95, 96, 97, 104, 105, 107, 120, 133, 134; **Satire I-5** lines 1, 2, 42; **Satire I-6** lines 1, 3, 6, 8, 25, 40, 45, 47, 48, 49; **Satire I-7** lines 33; **Satire I-9** lines 11, 16, 17, 19, 26, 27; **Satire I-10** lines 25, 26, 73, 84.

with Old Occitan syntax, and suggests that the vernacular language played a significant role in the teaching of Latin syntax to *pueri* or beginning students.⁴ At this very elementary level of learning Latin, the purpose of the syntactic glosses was, according to Reynolds, to reconfigure the text in the name of clarity and “the need to present a more familiar syntactic surface” (1990a: 33), one of a number of strategies—including lexical glosses—to relate the Latin to the vernacular (1990a: 47).

Reynolds analyzes a sample of 150 lines of the *Satires* (unfortunately these lines are not identified in Reynolds’ articles). She reports (1990a: 36) that the dominant word order in declarative matrix clauses is SVO: 50 out of 60 matrix clauses are SVO.⁵ “After conjunctions,” 27 out of 42 examples are SVO, while “the remaining fifteen feature inversion (11 VSO, 2 OVS, and 2 VOS).” All 12 interrogatives in Reynolds’ sample “place the verb immediately after the interrogative pronoun;” the 16 object relatives are always object relative pronoun – VS, while the 7 subject relatives all have VO order after the subject relative pronoun.

Reynolds appears to be unaware of the issue of V2 syntax, so she does not report the position of the verb in the 10 non-SVO matrix clauses, but the data she reports are broadly compatible with a V2 grammar. It is telling that she reports no instances of SOV. The incidence of nonsubject-initial matrix clauses is lower than what we report below, but it is likely that some of the 11 cases of VSO order after “conjunctions” involve clause-initial adverbial elements that trigger V2. Obligatory inversion in questions and in object relatives are both hallmarks of OOc grammar (Jensen 1986: 390). The two instances of VOS cited by Reynolds are a rare but attested order in OOc, found when the object is pronominal or the subject is heavy (Jensen 1986: 387).

With respect to VP-internal order, Reynolds finds that the direct object is consistently re-ordered to occur before the indirect object, perfectly in line with the attested facts of Old Occitan (Jensen 1986: 103).

Reynolds (1990b), in an addendum to her 1990a study, flags one instance of NP-internal order generated by the glosses as aberrant from the standpoint of OOc word order. In *Satires* 1.1, l. 55, *magno de flumine*, ‘from the big river’ the adjective is reordered to *de flumine magno*.⁶ Reynolds suggests that this is “a contravention of vernacular practice, where the equivalent adjective *grant* normally precedes the nouns” (1990a: 66). But noun-adjective order in OOc is quite free, as Reynolds acknowledges elsewhere (1990a: 39); see also Jensen (1986: 49). In the case of *grant* ‘large’, a standard dictionary of Old Occitan provides an example of postnominal

4 Reynolds contrasts this vernacular-influenced glossed word order with the idealized order of the *ordo naturalis* in medieval grammatical theory and reflected, she claims, in glossed texts aimed at more advanced learners. She does not, however, go so far as to say that the sequential glossing entirely reflects OOc syntax.

5 Exactly what Reynolds considers to be SVO is unclear; in particular is it unclear whether she excluded sentences lacking overt subjects or objects or adjunct topics. The proportion of SVO that she reports is high compared to our finding of 12/93 SVO matrix clauses, but this may be because she excluded tokens missing subject or object, or Adjunct-V-S order.

6 As is typically the case in BnF 7979, prepositions are glossed together with the following word; thus, in this instance the sequence glosses **i** on *magno* and **h** on *de flumine* to generate the order *de flumine magno*.

*grant: Antonis e lodist P. Quzol aviant paraulas **grandas** antre lor.*⁷ In the specific example cited by Reynolds, the placement of the adjective is contrastive, as *flumine* and *magno* are contrasted with *fonticulo* ‘little spring’.⁸ See (14b) below for an example of prenominal *magno*s in a non-contrastive context.

In sum, although Reynolds (1990a, 1990b) is unaware of the V2 syntax of OOC, her assessment that “features of the converted word order suggest a strong Occitan influence” is accurate. The one example where she suggests a departure from OOC word order is not unusual at all, in context.

5.2 Major constituent order in matrix clauses

Table 3 summarizes the placement of the verb in 94 glossed matrix clauses in BnF 7979, showing V2 order at 48%, followed by V1 at 32%, V3 at 13% and V4 at 1%. Comparing these proportions to quantitative descriptions of OOC corpora given in Sitaridou (2012) examining 12th and 13th century notary texts, and Wolfe (2018) drawing on the *Vida* of Saint Douceline, a late 13th century prose text, we find the glossed word orders from the Latin *Satires* of Horace to be strikingly similar to OOC prose texts.⁹

Table 3: A comparison of verb placement in matrix clauses

	<i>Satires</i> early 12 th c. Glossed order n=94	<i>Chartes</i> late 12 th c.–early 13 th c. Sitaridou (2012: 573) n=157	<i>Douceline</i> late 13 th c. Wolfe (2018: 3) n=622
V absent	5%	—	—
V (unique)	1%	—	—
V1	32%	12.5%	7.56%
V2	48%	65%	52.73%
V3	13%	23%	29.74%
V4	1%		8.04%
V5	0		1.29%
V6	0		0.64%

7 *Justice-Dienne*, 26v; Olivier 2009: 642:

(i) *Antonis e lodist P. Quzol aviant paraulas **grandas** antre lor*
 A. and the.said P. Q. had words big between them
 ‘A. and the aforementioned P. Q. had a violent exchange of words’

8 The original passage is *magno de flumine malle[m] quam ex hoc fonticulo tantundem sumere* ‘I had rather draw [so much] from a great river, than the very same quantity from this little fountain’. See example (12a).

9 What is counted as a preverbal constituent can vary from one study to another. For example, Sitaridou (2012) does not count clause initial negation as a constituent, while we do, given that it systematically allows pre-verbal clitics in OOC, consistent with the Tobler-Mussafia Law (Jensen 1986: 112; Mériz 1978: 157–8).

The proportion of V2 orders is comparable to both *Douceline* and the notary documents or *Chartes*, as is the proportion of V3+ orders at a total of 14%, although it is lower than the somewhat later Old Occitan data and includes no cases of V5+.

The proportion of V1 in the BnF 7979 glosses is, however, substantially higher than in the Old Occitan texts. It is here where we clearly see the constraints imposed by the Latin source text. These include, first, the absence of the discourse particle *si*, which is widely understood to indicate topic continuity and to serve as a last-resort expletive to satisfy V2 requirements in Old Occitan and other Old Romance varieties (see, for example Donaldson 2015; Wolfe 2018: 8–9; Salvesen 2013, 2020). If the glossator intended these sequences to closely approximate Old Occitan literary prose, then it is no stretch to imagine that a number of the V1 matrix clauses would have been preceded by *si*, rendering them V2. See section 5.4 for further discussion of V1 matrix clause order.

An additional factor contributing to the high incidence of V1 is the high proportion of null subjects in the source text. Fifteen, more than half of the twenty-nine V1 examples in our glossed data involve null subjects. As we saw in 3.1, V1 sentences with null *pro* discourse topics are a normal feature of OOc syntax. It is not clear whether Sitaridou and Wolfe included null subject data in their totals, but if these 15 *pro* V . . . examples are removed from the BnF 7979 data, the proportion of V1 tokens becomes even closer to their findings.

5.2.1 V3 and V4 configurations

13% of matrix clauses present a V3 order and 1% present the rare V4, shown in Table 1. These configurations remain remarkably loyal to the information structure in Horace’s text while respecting the constraints of a relaxed early Romance V2 grammar. Example (8) presents a typical example of V3. The glossed Latin text is followed by the reordered version, interpreted in terms of an articulated left periphery, outlined for Old Occitan in section 3.1.

- (8) Gloss: **a** **b** **d** **f** **h** **g** **i** **k** **c**
 Latin: *primum* *ego* *me* *illorum* *dederim* *quibus* *esse* *poetas* *excerpam*
e
numero
 Reordered:

 a **b** **c** **d** **e**
[FrameP *primum* [FocusP *ego* [FinP [Fin0 *excerpam*] *me* *numero*
 first I select.1SG.FUT me.ACC list.ABL
f
illorum . . .]]]
them.GEN

“First, I myself will take my name from the list of such . . .”

Satires I–4, 1.39–40 (BnF 7979 f.19r–1, 1.10–11)

The Latin sentence begins with a frame-setter, *primum*, followed by a focused element, the first-person pronoun *ego*, as the speaker says that he will be the first to

These two examples further demonstrate the glossator’s virtuosity in using the syntactic resources of Occitan to render the information structure of Horace’s verse.

5.2.3 Subjects

Given that OOC is a null-subject language, overt pronominal subjects are generally only introduced when the subject is contrastive or focal, as we noted in 3.3. Otherwise, non-lexical subjects are null. Under such conditions, overt pronominal subjects should predominantly occur pre-verbally. This is the case reported in Sitaridou (2012: 573)), where 26/28 (93%) pronominal subjects in her corpus are preverbal. The same holds for the glossed *Satires*; 11/12 (92%) pronominal subjects occur to the left of the verb. The one case in which it is post-verbal involves a V1 yes/no question.

The glossator’s systematic placement of subject pronouns to the left of the verb does not simply reflect preverbal position in Horace’s Latin. Our data show that the glossator actively follows the discourse constraint on the use of subject pronouns. In the following example from *Satires* I-2, lines 18–19 (BnF 7979, f.17v–1, l.7–8), the glossator reorders the subject demonstrative pronoun *hic* ‘he, this man’, which is post-verbal in the Horace text, so that it occurs in the second position of a V3 matrix clause. The finite verb *facit* ‘makes’ is underlined.

- (11) Gloss: **h m n n l k i**
 Latin: at in se pro quaestu sumptum facit **hic**
 Reordered:
h i k l m n n
at hic facit *sumptum* *in se* *pro quaestu*
 yet this.NOM make.3SG.PRES expense.ACC in himself.ABL for profit.ABL

“But surely he spends on himself in proportion to his gains.”

In contrast to pronominal subjects, nonspecific lexical subjects occur post-verbally 10/12 times. For example, *populus* ‘people’ Sat.I-1, l.66; *alius casus* ‘some mishap’ Sat.I-1, l.81; *quidam* ‘someone’ Sat.I-3, l.22 are preverbal in Horace’s Latin, but post-verbal in the grammar of the glossator. Since subjects have no privileged preverbal position in a V2 grammar (as they do in non-V2 SVO languages), they occur pre- or post-verbally according to their specificity properties (Diesing 1992) and their function in relation to the broader discourse. The overall proportion of pre-verbal vs post-verbal subjects is thus virtually equal in the BnF 7979 glosses: 26/45 matrix subjects are pre-verbal while the remaining 19 are post-verbal.

As we observed in 3.1, the post-verbal position of subjects in sentences with non-subject topics is a particular hallmark of V2 grammar. This hallmark is also found in the BnF 7979 glosses. Of V2 sentences containing an overt subject, verb, and at least one other overt constituent, 24 have SVO order, while 7 have XP-V-NP_{subject} order, where XP is the nonsubject topic.¹⁰ The resulting total of 30% is fully in line

¹⁰ One of the seven cases involves an XP-V-O-S sequence, where the object is a direct object pronoun and the subject is heavy. As we discuss below, a direct object pronoun may not occur clause finally unless it is right-adjacent to the verb. The heavy subject is ostensibly right dislocated.

with the percentage of XP-V-NP_{subject} data in present-day V2 languages (Lightfoot 1999: 156), reinforcing our claim that the glossator is following a V2 grammar. It is very difficult, in turn, to see how this word order could be the product of a didactic *ordo naturalis* based on principles such as “subject precedes verb”.¹¹

5.2.4 Objects

Objects are overwhelmingly post-verbal in matrix clauses. 93% (62/67) of objects occur post-verbally; 55% of these objects are pre-verbal in the Latin text. Of the five preverbal objects, three are straightforward nonsubject topics in a V2 syntax. One is OVS while the other two have the emphatic object pronoun *te ipsum* ‘you yourself’ in an OV sequence with a null subject. A fourth example involves quotative inversion, with the quotative complement preceding the verb and itself preceded by a left dislocated heavy subject to produce V3 order. The fifth example involves a proclitic object pronoun (example (14a) below).¹²

In nonfinite contexts involving infinitive forms, objects can occur both pre-verbally (12a) and post-verbally (12b) in the glossator’s reordered version of the Latin.

- (12) a. **e** **f** **g** **h** **i**
mallem *tantundem* *sumere* *de* *flumine* *magno*
 would.rather.1SG such.a.quantity take.INF from river.ABL big.ABL
- k** **l** **m**
quam ex *hoc* *fonticulo*
 than from this.ABL brook.ABL
 “I would rather have taken the quantity from a broad river than from this tiny brook” *Satires* I-1, lines 55–56 (BnF 7979, f.17r-1, l.22–23)
- b. **l** **m** **n** **o** **a** **b**
nolit *dare* *inopi* *amico* *quo* *possit*
 refuse.3SG.SUBJ give.INF poor.DAT friend.DAT which.ABL can
- c** **d** **e** **f** **g**
propellere *frigus* *que duram* *famam*
 drive.away.INF cold.ACC and harsh.F.S.ACC hunger.F.S.ACC
 “[he] would grudge a poor friend the wherewithal to banish cold and hunger’s pangs” *Satires* I-2, lines 5–6 (BnF 7979, f.17r-2, l.43–44)

We see that the glossator orders the verb and its object in a manner consistent with the grammars of both Old Occitan and Old French briefly described in section 3.4. In matrix clauses, when the verb is finite, lexical objects never occur in preverbal argument position in the BnF 7979 glosses, although this sequence is naturally frequent

11 A good description of the *ordo naturalis* as taught in 10th century St. Gall is given by Grotans 2006: 164–7. As Grotans observes, Priscian, the most influential source for medieval grammarians, lacked terminology other than part-of-speech labels for addressing word order (e.g., “the noun must precede the verb”). Other classical precursors supplied the more useful notion, still descriptively inadequate for a V2 language, that the subject should precede the predicate.

12 *Satire* I-2, l.44; *Satire* I-3, l.34; *Satire* I-9, l.25; *Satire* I-2, l.35–36; *Satire* I-1, l.66.

in Horace’s Latin. (13) is a representative example of how the glossator reorders SOV in the source text.¹³

(13) Gloss: **h** **b** **d** **c** **a**
 Latin: ecce crispinus minimo me provocat
 behold Crispinus.NOM long.odds.ABL me.ACC challenges

Reordered:

h **a** **b** **c** **d**
Ecce provocat Crispinus me minimo
 “See, Crispinus challenges me at long odds”

Satires I–4, 1.14; BnF 7979 f.18v–2, 1.33

Such categoric avoidance of the SOV order in the source text is exactly what we expect in a V2 grammar of the Old Romance kind. In a finite clause, the only way to generate SOV order would be a V3 configuration with an overt topic subject and focused object. Given the glossator’s fidelity to OOc grammar and the information structure of the source text, there is no reason to expect such a marked configuration. Table 4 summarizes the placement of subjects versus objects in the BnF 7979 glosses.

Table 4: Subject-object asymmetry in matrix clauses

	subject		object	
pre-verbal	26	58%	5	7%
post-verbal	19	42%	62	93%
Total	45		67	

This brings us to the question of nonlexical objects, specifically object pronouns. Recall that in OOc, object pronouns come in both strong and clitic forms (section 3.3), where the latter can be proclitic or enclitic on the verb. In the BnF 7979 glossed data, both oblique and accusative pronouns are predominantly post-verbal but always adjacent to the verb if they are the sole argument. What would be a direct object pronoun in OOc is only ever separated from the verb when it is one of two internal arguments, or when it occurs within a small clause complement. In some cases, such as (14c–d), the object pronouns would clearly have to be interpreted as strong forms, or, in the case of obliques, a preposition plus strong pronoun. Below are examples of what we believe to be a proclitic (14a) and an enclitic (14b), a pronominal DO + IO combination (14c), and a pronoun inside a small clause (14d), showing only the reordered data.

(14) a. **a** **b** **c** **d** **e**
 at ipse **mihi** plaudo domi
 but self.M.NOM me.DAT applaud.1SG.PRES house.LOC
 “but I applaud myself at home” *Satires* I–1, 1.66; BnF 7979 f.17r–1, 1.33

¹³ *Ecce* immediately precedes a line break in the manuscript; the gloss h over *ecce* is part of the gloss sequence of that line, indicating that it should be read in the final position of that line, which happens to be the beginning of the sentence in (13).

- b. **k o l m n o**
 brute oro **te** per magnos deos
 Brutus.VOC implore.1SG you.ACC by great.PL.ACC god.PL.ACC
 “O Brutus, by the great gods I implore you”
Satires I-7, 1.33; BnF 7979 f.20v-2, 1.37
- c. **m acos n o p q**
 insuevit optimus pater **me hoc**
 accustomed.3SG best.NOM father.NOM me.ACC this.ABL
 “The best of fathers taught me this (habit)”
Satires I-4, 1.105; BnF 7979 f.19r-2, 1.26
- d. **a b c d e**
 non vult uxor [te salvum]
 NEG want wife.NOM you.ACC well.ACC
 “Your wife does not want you well”
Satires I-1, 1.84; BnF 7979 f.17r-2, 1.1

The contrast between proclitic (14a) and enclitic (14b) conforms to the conditions for proclisis vs enclisis in OOc identified by Donaldson (2016), mentioned in 3.3. In (14b), the overt preverbal constituent is a vocative, which we would expect to have the status of a left dislocated element. This is exactly the condition for enclisis according to Donaldson. In (14a), the specifier immediately to the left of the verb is occupied by the overt subject. This satisfies Donaldson’s condition for proclisis.

In (14c), equally interestingly, the glossator only reorders two lemmata: the two object pronouns. This would produce DO-IO word order, assuming the equivalent verb in Old Occitan: *acostumar*.¹⁴ This reordering also conforms to the glossator’s systematic avoidance of placing the direct object clause finally unless it is right-adjacent to the verb. In (14d), the glossator radically reorders the Latin lemma *non uxor salvum te vult*. In OOc, the small clause could easily be construed as a finite subordinate clause along the lines of *que tu sias san* ‘that you be well’. In any event, these appear to be the only patterns in our sample where accusative object pronouns are separated from the verb.

Finally, we remark that in our examination of all object pronouns in the first book of *Satires*, there is not one violation of the Tobler-Mussafia law. That is to say, we find no re-orderings whereby a simple object pronoun occurs clause initially.

5.3 Major constituent order in embedded clauses

We saw in 3.2 that OOc has an asymmetrical V2 grammar, with relaxed V2 in matrix clauses and frequent nonsubject-initial word order in matrix clauses, but overwhelmingly SVO order in embedded clauses. Wolfe’s (2018) data on embedded clauses is taken from clausal complements introduced by *que*. The closest Latin counterpart to *que* complements are *ut* ‘that’ complement clauses. In our sample of the glossed *Satires* we find three of these having an overtly expressed subject. Two show *ut* SV word order (15a). The third involves an impersonal construction with V1 order (15b).

¹⁴ The lexical gloss *acos* written above *insuevit* further supports such an analysis.

- (15) a. **a b c a b c d e**
quin etiam illud accidit ut quidam demeteret testes
 even nay that.NOM befalls that someone.NOM down.put testicles.ACC
 “Nay, it even happened that someone sheared off the testicles”
Satires I-2, 1.44-45, BnF 7979 f.17v-1, 1.33-34
- b. **a b c a c d**
est opere-pretium audire [. . .] ut-que cadat illis
 is effort.GEN-worth.NOM hear.INF that-and falls.SUBJ them.DAT
e f g h
voluptas corrupta multo dolere
 pleasure.NOM perverted.NOM much.ABL pain.ABL
 “It is worthwhile to hear [. . .] that perverse pleasure befalls them with
 much pain”
Satires I-2, 1.37-40, BnF 7979 f.17v-1, 1.26-28

There is no evidence that the verb raises any higher than T in any of these examples, although V1 in the impersonal pattern in (15b) is compatible with placement of the verb in a higher position. Wolfe (2018: 9, ex.24a) reports embedded V2 in OOC complements of bridge verbs such as *ausir/auzire* ‘hear/listen’, the cognate of the matrix verb *audire* in (15b), as in (16):

- (16) *Vos avez ben, [som cug,] auzit [que-l rei pot aver per marit].*
 You have well I think heard that-the king can have for husband
 “You have well, I think, heard that the king, she can have for a husband.”
Flamenca, 49-50

The ten accusative + infinitive (ACI) constructions in our data also present evidence for the structure of OOC embedded clauses, albeit indirect. Since Latin ACIs find their clearest equivalents in finite complement clauses in the vernacular, we compare the reordered elements of the ACI with the *ut* ‘that’ clauses discussed above, under the assumption that the accusative argument is construed as the subject in the vernacular and the infinitive as the finite verb. It turns out that they pattern exactly alike. 8/10 of ACI complements occur with overt subjects; 5/8 have SV order as in (17a). The remaining three have VS order, but all of the latter involve unaccusative infinitival verbs, as in (17b), analyzable as having what would be the finite verb in T and the subject in situ. Again, there is no evidence of embedded V2. (17) shows only the reordered version of the text.

- (17) a. **d e f g h k**
ne putes me compilasse scrinia crispini
 so.that.not think.2SG me.ACC to.have.plundered rolls.ACC Crispinus.GEN
 “so that you will not think that I have rifled the rolls of Crispinus”
Satires I-1, 1.120-121, BnF 7979 f.17r-2, 1.37-38
- b. **a b c d e**
est equum descendere iuvenes huc
 is right to.go.down young.ACC hither
 “it is well that young men come down hither”
Satires I-2, 1.34, BnF 7979 f.17v-1, 1.23

As we noted in 5.1, Jensen (1986: §1122) observes that in OOC nonsubject relatives, the verb is frequently in second position after the relative pronoun. We saw that Reynolds found this without exception for both relatives and *wh*-questions in her sample of the glossed *Satires*. We find the same for nonsubject relatives as shown in (18), the reordered version of the text.

- (18) **g h i**
quem inducit fabula
 whom describes story.NOM
 “whom the story describes” *Satires* I–2, l.21–22, BnF 7979 f.17v–1, l.10–11

Further evidence of embedded V2 syntax is found in embedded adjunct clauses. 11/28 such clauses present subject-verb inversion (that is, V2), occurring, for example, in all of the 5 *ubi* ‘when’ clauses in our data.

- (19) **a b c**
Ubi pulsat consultor
 when knocks client.NOM
 “when a client comes knocking” *Satires* I–1, l.10–11, BnF 7979 f.16v–2, l.25

The one example in our data of an indirect question also shows subject-verb inversion.

- (20) **d e f g h i k**
cur stringat malus ingrata ingluvie preclaram rem
 why strips wicked.NOM thankless.ABL gluttony.ABL famous.ACC thing.ACC
 “why this wicked one, in his thankless gluttony, strips the noble estate [of his grandsire ...]” *Satires* I–2, l.7–8, BnF 7979 f.17r–2, l.45–46

Here again the glossed data converge with the Old Occitan data showing inversion, that is, V2, in relatives and questions. Inversion is harder to demonstrate in subject relatives but we present evidence for it as well in section 5.5.

Table 5: Embedded clause types and their glossed word order

	Total	Overt Subj.		Null subj.	
		SV	VS	V-XP	XP-V
Tensed complement clauses (<i>ut</i> ‘that’)	7	2	1	4	0
Non-finite complement clauses (ACI)	10	5	3	2	0
Object relative clauses	6	0	5	1	0
Adjunct clauses (<i>ubi</i> , <i>cum</i> ‘when’; <i>si</i> ‘if’; <i>nam</i> ‘for’, etc.)	28	4 ¹⁵	11	13	0
Indirect questions	1	0	1	0	0
Subject relative clauses	7	7	0	0	0
Total embedded clauses	59				

¹⁵ One of these tokens involves a *quanto... tanto* ‘as much as’ clause pair where the verb is final preceded by a comparative adjective in both cases.

Embedded clauses show two clear tendencies: while both types of declarative complement clauses (*ut* and ACI clauses) show no evidence of V raising to the left periphery, everything else—nonsubject relatives, adjunct clauses, and indirect questions—provide robust examples of V to C raising, facts entirely in line with what is reported for OOC.

5.3.1 Germanic inversion

The BnF 7979 glosses attest so-called Germanic inversion, where the tensed auxiliary is separated from the participle and placed in second position:

- (21) Gloss: **a h i b g f d e c**
 Latin: ille datis vadibus qui rure extractus in urbe¹⁶ est.
 Reordered:
Ille qui est in urbe extractus rure datis vadibus
 3SG.M who is in city.ABL dragged.NOM country.ABL given.ABL bail.ABL
 “He who is dragged from the country into the city, having given surety,”
Satire I.1, 1.11 (BnF 7979, 16v–2, 1.26)

Needless to say, Germanic inversion is a characteristic V2 property. The example in (21) also attests to the specific OOC property of inversion (that is, V raising) in relative clauses noted by Jensen, which we turn to in the next section.

5.3.2 Relative clause internal order

We have already seen that the BnF 7979 glosses show the inversion pattern in relative clauses noted by Jensen (1986: 390), as confirmed by Reynolds. In OOC object, or more generally nonsubject relatives, this involves the verb being placed immediately after the relative pronoun, as in Jensen’s example (22):¹⁷

- (22) *belha domna, en cui renha senz e beutatz*
 beautiful lady in whom dwell reason and beauty
 “fair lady, in whom reason and beauty dwell”
 (*Peirol* XII 36; Jensen 1986: 390)

Jensen’s examples of inversion are limited to object relatives and a *when* clause (1986: 390), but Germanic inversion in the subject relative in (21) suggests that the tensed verb may raise in subject relatives as well. General inversion in relative clauses in turn indicates that OOC relatives are not subject to the doubly filled Comp filter: that

16 H.S. Fairclough’s edition of the *Satires* has, with no noted exceptions, *in urbem*, giving the directional reading ‘into the city’. Our copyist’s *in urbe*, suggesting a locative rather than a directional reading, might be a scribal error, but if in fact the glossator interpreted the phrase as “who is in the city (having been) dragged from the country”, our argument for Germanic inversion here is weakened.

17 Jensen observes that inversion (V2) in nonsubject relatives is not obligatory: he cites an example from the same text with the same verb and pied-piped relative pronoun *en cui* ‘in whom’ which does not trigger inversion. Jensen does not speculate on what the conditions for this alternation are, but a cursory inspection of Jensen’s examples suggests that inversion may occur in nonrestrictive relatives, noninversion in restrictives.

is, relative pronouns occupy Spec, CP, and the raised verb is free to move to Comp. This supposition is consistent with another very striking word order property of OOc relative clauses observed by Fanselow and Féry (2006). They provide the following undated example:

- (23) *la justícia que grant áig a mandar*
 the legal.power which great have.1sg to dispose
 “the great legal power which I have at my disposal”
 (Fanselow and Féry 2006: 7, fn. 8)

The interesting thing about this example is that the adjective *grant* ‘great’, which modifies the head noun *justícia*, occurs after the relative pronoun, within the relative clause. Fanselow and Féry analyze this as a discontinuous NP. But under a head raising analysis of relative clauses of the sort proposed by Kayne (1994), (23) could be analysed as stranding of the head NP by raising of the N⁰ head *justícia*. Under this analysis, the head NP consists of the relative pronoun plus the head noun and modifiers. This head NP is first raised to the specifier of the relative CP; (23) would then be derived by raising the head N⁰ further. Regardless of the exact derivation, the relevant fact here is that this crosslinguistically highly marked order is attested in the free relative in the second V1 sentence in (10), repeated in (24) below with the original Latin:

- (24) Gloss: e f e d b c a
 alia ambulat
 Latin: est qui inguen ad obscenum subductis usque facetus ;
 Reordered:
Est qui facetus subductis usque ad obscenum inguen
 is who elegant.NOM pull.up.ABL even to obscene.ACC groin.ACC
 “There is the elegant one who pulls [his toga] up even to his obscene groin.”

This example shows Horace’s original Latin elaborately rearranged to derive an order found in few languages other than Old Occitan.

5.4 Summary

In the analysis of the BnF 7979 glossed data above, we have found a number of properties typical of an earlier Romance relaxed V2 grammar as well as properties specific to OOc. The former include a plurality of V2 tokens; the existence, in smaller numbers, of V1, V3, and V4; SV... order in complement clauses as opposed to XP V... order in root clauses with a smaller statistical preponderance for subject-initial order in the latter; V1 clauses in rhematic and topic continuity contexts; and Germanic inversion. Specific OOc properties include inversion (V to Comp) in relative clauses and coexistence of proclitic and enclitic object pronouns, distinguished in specific discourse-pragmatic contexts; and the unusual adjective-relative clause order in (23).

In discussing to what extent the *ordo naturalis* (or other normative medieval grammatical precepts) might have dictated the word order generated by the BnF 7979 glosses, Reynolds writes, “*Ordo naturalis* was a long enshrined grammatical tool with

which to unravel the rhetorical intricacies of Classical syntax. Yet despite its expositional pedigree it was not always adhered to by the [BnF 7979]P glossator” (1990a: 38). We see no evidence that the *ordo naturalis* played any role in the BnF 7979 glosses at all. The main precept referred to by Reynolds, avoidance of discontinuous constituents, is a general byproduct of the transition from Latin to early Romance languages. Other hazy dicta of the *ordo*, such that “agents” should precede “actions”, are directly contradicted by glossings such as (10), where subject-predicate order in Horace is reversed by the rhematic V1 order in the glosses in three successive sentences.

The glossator of BnF 7979 was a highly skilled interpreter of Horace’s Latin. His or her objective was to transform Horace’s text into a fluent vernacular syntax with plausible discourse-pragmatic cohesion, within the constraints of the lexicon and information structure supplied by Horace.

6 How the text was read

Linguists reading to this point may reasonably ask: Accepting that the sequence glosses in BnF 7979 follow Old Occitan word order, how was the text actually read? Reynolds and other scholars of medieval glossed texts have generally held that the glossing was for pedagogical purposes: The glossator-teacher added the sequence glosses, the lexical glosses, and the neumatic notations for her or his use in lecture or instruction. But if part of this instruction was reading the text in OOC word order, is it plausible to imagine that the instructor read the words of Horace using Occitan grammar?

The answer is yes. The answer taps into another longstanding debate, this one among Romanists, extending back to Wright (1982). Wright’s view is that prior to the 8th century Carolingian reforms in orthography and reading practice, Romance speakers did not distinguish between Latin and the incipient Romance vernaculars; standard procedure was to write in what a modern reader would recognize as “Latin”, but to read (meaning, in particular, read aloud) in the local vernacular. This would mean, according to Wright, not simply reading the Latin orthography according to contemporary local Romance phonology, but in some cases employing what Wright calls “logographic” reading/writing: substituting a current Romance vernacular term for a Latin item that had lapsed into disuse, e.g., *feu/fuego/foc* for *ignis*.

Some aspects of Wright’s thesis remain highly controversial, in particular the role of the Carolingian reforms and the impact of specific individuals such as Alcuin on those reforms, as well as the scope of “logographic” reading/writing. But there is no doubt that there are other early Romance glossed texts with syntactic glossing that attest to the practice of what we have called vernacular reading. A well-studied example nearly contemporary with BnF 7979 are the *Glosas Emilianenses*, glosses inserted in the 11th century in the Latin sermons collected in the *Códice Emilianense* 60, part of a collection of religious writings copied several centuries earlier. The glosses are the subject of a recent doctoral dissertation (Hagemann 2009) and much discussion by specialists, including Wright (1982), who holds that the purpose of the glosses was “to facilitate intelligible oral delivery”. The glosses include alphabetic sequence glosses much like those we have discussed in BnF 7979, grammatical glosses,

and lexical glosses in Latin, an early Romance variety (interpreted by most scholars as Old Spanish or a precursor of it), and in two cases, Basque. There is no serious dispute regarding the fact that the syntax indicated by the glosses is early Romance, not Latin. There is some dispute over whether the purpose of the glossing is for classroom use, or, as Wright argues, oral performance of the text. As Hagemann points out, the fact that the source texts are sermons, not normally material used in the monastic classroom, supports the latter view.

To conclude this section, we cannot be sure of how the BnF 7979 Horace was read in every detail, but the existence of contemporary Romance texts glossed in an almost identical way for oral delivery suggests that it too was read out according to the OOc syntax we have demonstrated in this paper. We cannot be sure whether “logographic” lexical substitution (for example, when an Occitan lexical gloss is present in the manuscript) took place or not, but numerous features of the orthography, for example the spelling of the proper name *Maltinus* as <*Malchinus*>, indicating *t*-palatalization, indicate that it was read with OOc phonology.

7 Conclusion and further prospects

In this paper, we have argued that Latin texts of early Romance provenance with syntactic glossing can reflect, rather accurately, contemporary Romance syntax, and further suggested that they might even serve as a source for historical syntax research. The first claim is not original here, although we believe we have established the case much more strongly than previous research for the particular manuscript under discussion. The second claim is a promissory note: until historical syntacticians and other specialists gain confidence in the vernacular character of the glossed material, more extensive documentation is required. But we believe we have shown, for example in the data on relative clause-internal syntax in 5.6, that even very intricate features of contemporary syntax can be identified in these texts.

References

- Bauer, Bernhard, Rijcklof Hofman, and Pádraic Moran. 2017. St Gall Priscian Glosses, version 2.0. URL www.stgallpriscian.ie/resource, [accessed 24 June 2020].
- Benincà, Paola, and Nicola Munaro, ed. 2010. *Mapping the Left Periphery: The Cartography of Syntactic Structures*, vol. 5. Oxford: Oxford University Press.
- Benincà, Paola, and Celia Poletto. 2004. Topic, Focus, and V2. In *The Structure of CP and IP*, ed. Luigi Rizzi, 52–75. Oxford: Oxford University Press.
- Biberauer, Theresa, and Ian Roberts. 2005. Changing EPP parameters in the history of English: Accounting for variation and change. *English Language and Linguistics* 9:5–46.
- Blom, Alderik H. 2017. *Glossing the Psalms: The Emergence of the Written Vernaculars in Western Europe from the Seventh to the Twelfth Centuries*. Berlin: De Gruyter.
- Diesing, Molly. 1992. *Indefinites*. Cambridge, MA: MIT Press.
- Donaldson, Brian. 2015. Discourse functions of subject left dislocation in Old Occitan. *Journal of Historical Pragmatics* 16:159–86.
- Donaldson, Brian. 2016. Preverbal subjects, information structure, and object clitic position in Old Occitan. *Journal of Linguistics* 52:37–69.

- Draak, Maartje. 1967. Construe marks in Hiberno-Latin manuscripts. *Mededelingen der Koninklijke Nederlandse Akademie van Wetenschappen* 20:261–282.
- Draheim, Joachim, and Günther Wille. 1985. *Horaz-Vertonungen vom Mittelalter bis zur Gegenwart: Eine Anthologie*. Amsterdam: B. R. Grüner.
- Fairclough, H. Rushton. 1942. *Horace: Satires, Epistles and Ars Poetica. With an English Translation by H. Rushton Fairclough*. Loeb Classical Library. Cambridge, MA: Harvard University Press.
- Fanselow, Gisbert, and Caroline Féry. 2006. Prosodic and morphosyntactic aspects of discontinuous noun phrases: A comparative perspective. Ms., University of Potsdam.
- Grotans, Anna. 2006. *Reading in Medieval St. Gall*. Cambridge: Cambridge University Press.
- Hagemann, Kristin. 2009. The Emendation of a Manuscript: The Purpose of the *Glosas Emilianenses*. Doctoral Dissertation, University of Oslo.
- Hale, Mark. 1987. Studies in the Comparative Syntax of the Oldest Indo-Iranian Languages. Doctoral Dissertation, Harvard University.
- Jensen, Frede. 1986. *The Syntax of Medieval Occitan*. Tübingen: Niemeyer.
- Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge, Mass: MIT Press.
- Korhammer, Michael. 1980. Mittelalterliche Konstruktionshilfen und Altenglische Wortstellung. *Scriptorium* 34:18–58.
- Kosukegawa, Teiji, and John Whitman. 2018. On the significance of the glosses in Vietnamese Classical Chinese texts. *Journal of Vietnamese Studies* 13:28–49.
- Lemoine, Louis. 1994. Signes de construction syntaxique dans les manuscrits bretons du haut Moyen age. *Archivum Latinitatis Mediiævi* 52:77–108.
- Lightfoot, David. 1999. *The Development of Language: Acquisition, Change and Evolution*. Oxford: Blackwell.
- Marchello-Nizia, Christiane. 1995. *L'évolution du français: Ordre des mots, démonstratifs, accent tonique*. Paris: Armand Colin.
- Méris, Diana Theresa. 1978. Observations on object pronoun collocation with finite verb-parts in Medieval Occitan (to 1300). *Romania* 99(394):145–182.
- Munk Olsen, Birger. 1982. *Catalogue des manuscrits classiques latins copiés du IXe au XIIIe siècle, tome 2*. Paris: CNRS.
- Munk Olsen, Birger. 1996. Medieval manuscripts of the Latin classics: Production and use. In *Proceedings of the Seminar in the History of the Book to 1500, Leiden, 1993*, ed. Claudine A. Chavannes-Mazel and Margaret M. Smith. Los Altos Hills, CA: Anderson-Lovelace.
- Munk Olsen, Birger. 2007. Chroniques des manuscrits classiques latins (IXe-XIIes). VI. *Revue d'Histoire des Textes* 2:49–106.
- Munk Olsen, Birger. 2009. *L'étude des auteurs classiques latins aux XIe et XIIe siècles. Tome IV. 1, La réception de la littérature classique: travaux philologiques par Birger Munk Olsen*. Paris: CNRS.
- Nam, Phunghyun. 2010. Hankwukö-sa e issö ūi kukyöl charyo [Kukyöl materials in the context of the history of Korean]. *Kukyöl yönku* 124:5–36.
- Nilsen, Christine Meklenborg. 2020. Resumptive structures in a Gallo-Romance perspective. In *Variation and Change in Gallo-Romance Grammar*, ed. Sam Wolfe and Martin Maiden, 41–70. Oxford: Oxford University Press.
- Olivier, Philippe. 2009. *Dictionnaire d'ancien Occitan auvergnat*. Tübingen: Niemeyer.
- O'Neill, Patrick. 1992. Syntactical glosses in the Lambeth Psalter and the reading of the Old English interlinear translation as sentences. *Scriptorium* 46:250–256.
- Reynolds, Suzanne. 1990a. Ad Auctorum expositionem: Syntactic theory and interpretative practice in the twelfth century. *Histoire Epistemologie Langage* 12:31–51.

- Reynolds, Suzanne. 1990b. Ad Auctorum expositionem: Syntactic glosses on literary texts in the twelfth century. *Histoire Epistemologie Langage* 12:65–67.
- Reynolds, Suzanne. 1996. *Medieval Reading, Grammar, Rhetoric and the Classical Text*. Cambridge: Cambridge University Press.
- Reynolds, Suzanne. 2000. Inventing authority: Glossing, literacy and the classical text. In *Prestige, Authority and Power in Late Medieval Manuscripts and Texts*, ed. Felicity Riddy, 7–16. Woodbridge: York Medieval Press.
- Riou, Yves-François. 1990. Codicologie et notation neumatique. *Cahiers de Civilisation Médiévale* 131:255–280.
- Riou, Yves-François. 1991. Chronologie et provenance des manuscrits latins neumes. *Revue d'Histoire des Textes* 21:77–113.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of Grammar: A handbook of Generative Syntax*, ed. Liliane Haegeman, 281–337. Dordrecht: Kluwer.
- Robinson, Frank. 1973. Syntactical glosses in Latin manuscripts of Anglo-Saxon provenance. *Speculum* 48:443–475.
- Salvesen, Christine Meklenborg. 2013. Topics and the left periphery: A comparison of Old French and Modern Germanic. In *In Search of Universal Grammar: From Old Norse to Zoque*, ed. Terje Lohndal, 131–174. Amsterdam: Benjamins.
- Sitaridou, Ioanna. 2004. A corpus-based study of null subjects in Old French and Occitan. In *Romance Corpus Linguistics II: Corpora and Diachronic Linguistics*, ed. Claus D. Pusch, Johannes Kabatek, and Wolfgang Raible, 359–374. Tübingen: Narr.
- Sitaridou, Ioanna. 2012. A comparative study of word order in Old Romance. *Folia Linguistica* 46:553–604.
- Vance, Barbara, Brian Donaldson, and B. Devan Steiner. 2009. V2 loss in Old French and Old Occitan: The role of fronted clauses. In *Romance Linguistics 2009: Selected Papers from the 39th Linguistic Symposium on Romance Languages (LSRL), Tuscon, Arizona*, ed. Sonia Colina, Antxon Olarrea, and Ana Maria Carvalho, 301–320. Amsterdam: Benjamins.
- Sneyders de Vogel, Kornelis. 1927. *Syntaxe historique du français*. 2nd edition. Groningen–La Haye: Wolters.
- Wallenberg, Joel. 2009. Antisymmetry and Conservation of C-Command: Scrambling and Phrase Structure in Synchronic and Diachronic Perspective. Doctoral Dissertation, University of Pennsylvania.
- Whitman, John, Valerio Luigi Alberizzi, Masayuki Tsukimoto, Teiji Kosukegawa, Tomokazu Takada, Miyoung Oh, and Jiwon Park. 2010. Toward an international vocabulary for research on vernacular reading of Chinese texts (Hanwen Xundu). *Scripta* 2:61–83.
- Wieland, Gernot R. 1983. *The Latin Glosses on Arator and Prudentius in Cambridge University Library MS Gg. 5.35*. Toronto: Pontifical Institute of Medieval Studies.
- Wolfe, Sam. 2015. Verb-initial orders in Old Romance: A comparative account. *Revue Romaine de Linguistique* 60(2–3):148–172.
- Wolfe, Sam. 2016. On the left periphery of V2 languages: Evidence from Romance FIN and FORCE V2 systems. *Rivista di Grammatica Generativa* 38:287–310. Selected Papers from the 41st Incontro di Grammatica Generativa.
- Wolfe, Sam. 2018. Occitan, verb second and the Medieval Romance word order debate. In *Romance Languages and Linguistic Theory 13: Selected Papers from Going Romance 29, Nijmegen*, ed. Janine Berns, Haike Jacobs, and Dominique Nouveau, 315–336. Amsterdam: Benjamins.
- Wolfe, Sam. 2019. Redefining the typology of V2 languages: The view from Medieval Romance and beyond. *Linguistic Variation* 19(1):16–46. Special Issue ‘A Micro-Perspective on V2 in Romance and Germanic’.

- Wright, Roger. 1982. *Late Latin and Early Romance in Spain and Carolingian France*.
Liverpool: Francis Cairns.
- Zaring, Laurie. 2010. Changing from OV to VO: More evidence from Old French. *Janua:
Revista Philologica Romanica* 10:1–18.
- Zaring, Laurie. 2011. On the nature of OV and VO order in Old French. *Lingua* 121:1831–
1852.