

Asymmetry as a general cue for V2 (loss)

Pierre Larrivé

Université de Caen Normandie
pierre.larrivee@unicaen.fr

Cecilia Poletto

Goethe Universität Frankfurt am Main & Università degli Studi di Padova
poletto@em.uni-frankfurt.de

Francesco Pinzin

Goethe Universität Frankfurt am Main
pinzin@em.uni-frankfurt.de

Mathieu Goux

Université de Caen Normandie
mathieu.goux@unicaen.fr



How to cite: Larrivé, Pierre, Poletto, Cecilia, Pinzin, Francesco & Mathieu Goux. 2024. Asymmetry as a general cue for V2 (loss) Larrivé, Pierre, Poletto, Cecilia, Pinzin, Francesco & Mathieu Goux. 2024. Asymmetry as a general cue for V2 (loss). RLLT23, eds Lisa Brunetti, Ioana Chitoran & Alexandru Mardale. Special Issue of *Isogloss. Open Journal of Romance Linguistics* 10(7)/11, 1-22.
DOI: <https://doi.org/10.5565/rev/isogloss.410>

Abstract

This paper identifies a micro-cue correlating to verb second word order (V2) in two closely related Medieval Romance languages. As V2 is asymmetrically distributed in main rather than subordinate clauses, an asymmetry would be expected in phenomena assumed to relate to V2, such as subject inversion, null subject and enclisis. The loss of that asymmetry should therefore indicate the loss of the V2 word order rule. These assumptions are tested here by a quantitative analysis of a treebank of calibrated data covering the crucial period of change (from the 14th to the 16th century) for Medieval French and Venetian. The hard quantitative evidence provided demonstrates that the main versus embedded asymmetry is indeed a micro-cue of V2 structure, and of its loss in one of the two investigated languages.

Keywords: V2, French, Venetian, treebanks, main clause, subordinate, subject inversion, null subject, enclisis.

1. Introduction

A typologically rare option regarding clausal word order is that of the verb-second rule (V2), by which the inflected verb rises to the Left Periphery of the clause where it is preceded by some initial phrase (see Holmberg 2015 and Woods & Wolfe 2020 for a review). Found in most contemporary Germanic languages (except English), it is widely assumed that V2 was operative in at least some of the Medieval Romance languages, where it has been lost. Among these Medieval French (Le Coultre 1875, Thurneysen 1892, Foulet 1919, Adams 1987, Roberts 1993, Vance 1997, Labelle 2007, Mathieu 2013, Wolfe 2018, Larrivé 2019) and Medieval Venetian (Benincà 1983/1984, 1995, 2006, 2013, Poletto 2020a, Singh 2021) have been identified as V2 languages. The extensive research on V2 in these languages has however not yet fully explained the reasons for the loss of the V2 word order. One reason for this situation is purely empirical. Even looking at the most influential recent monographs (Meisel et al. 2013, Poletto 2014, Wolfe 2018), the putative V2 status of Medieval Ibero-Romance, Gallo-Romance and Italo-Romance has been mainly assessed based on representative examples from verse and prose literary texts from various regions. When systematic studies of a coherent empirical sets have been performed, they have generally relied on extended passage of sources rather than whole sources for reasons of expediency. Another reason is that, while the field is moving towards accepting that word-order change may be caused by a variety of factors (Poletto 2019, Wolfe in press), identifying and quantifying these cues is a delicate affair.

This paper investigates a scenario for V2 loss. We propose that the structure, acquisition, and evolution of V2 word order is correlated to micro-cues found in a grammar. One such cue is proposed here and analyzed systematically, which is the distributional asymmetry between main and subordinate clauses. This asymmetry is believed to characterize the distribution of V2 in most languages (*pace* Yiddish and Icelandic; see Diesing, 1990; Santorini, 1992 for Yiddish, Rögnvaldsson & Thráinsson 1990 for Icelandic, but see Walkden & Booth 2020 for data against the lack of asymmetry in both languages). That is because V2 derives from movement of the verb to the left periphery, and the left periphery is fully accessible only in main clauses due to the presence of the complementizer in embedded clauses. Several phenomena thought to be dependent on V2 such as subject inversion, null subjects, and enclisis, are believed to be similarly asymmetric. The expectation is that if the asymmetrical distribution of V2 is lost through time as a consequence of losing V2 itself, this loss should be reflected in all V2-dependent phenomena.

The purpose of this paper is to evaluate the correlation between V2 word order, its asymmetric distribution through time, and that of V2-dependent phenomena. This is achieved by a comparative quantitative analysis of treebank data made up of strictly comparable textual material. The comparison is conducted in two Medieval Romance varieties, French and Venetian, that seem to pattern closely as to V2 and its loss (e.g., Wolfe 2018 and references therein). Although these two medieval Romance languages are the focus of our investigation, the methodological and conceptual perspective articulated here can be applied to future comparative analyses.

The paper is organized as follows. The next section presents the methodological choices regarding the treebank investigation. Section 3 then reports on the distributional asymmetry of V2 through time in each language, and that of subject inversion, null subject and enclisis. The significance of the result is summarized by the conclusion (Section 4), that presents directions for future investigations.

2. The study: objectives and methodology

Our objective is to assess the potential correlation between V2 change – measured as relative frequency of V2 linear word order in main clauses – and asymmetric distribution of phenomena believed to relate to it. If asymmetric distribution is a cue for the structure, acquisition, and evolution of V2, the expectation is that there should be a correlation between the evolution of asymmetric distribution of V2 and that of V2 dependent phenomena through time.

First, we evaluate V2 itself. To this end, the proportion of main clauses linear V2 through time is compared to the proportion of V1 on the one hand and V3 and V4 on the other. The specific methodological challenges in how to count these occurrences are addressed at the end of this section.

As a first dependent phenomenon, the share of subject inversion in main and subordinate is established through time, as it is considered prime evidence of the involvement of a V2 word order rule (Adams 1987, Roberts 1993, Vance 1997 on OFr, Benincà 2006, Poletto 2014 on OFI and other old Italo-Romance varieties; but see Kaiser 2002, Rinke 2007, Rinke & Meisel 2009 and Kaiser & Zimmermann 2011). In (1)-(2) we observe how both languages present subject inversion, also between the finite verb and the past participle/infinitive (the so-called “Germanic inversion”).¹ Although this phenomenon is well attested in both languages, the question is how often this happens, whether it happens more in main than in subordinate clauses, and to what degree.

- (1) Old Venetian, Lio Mazor 15 (1311)
 ... e così avev' e' lo me' cortel da ferir en man trato.
 ... and so had.1SG I the my knife to stab.INF in hand taken
 ‘And so, I pulled out in my hand my dagger.’
- (2) Old French, Atiremens (1314)
 Et si doit le serjant escharir le serement
 and so must.3SG the sergeant make.pronounce.INF the pledge
 ‘And so, the sergeant should have the pledge pronounced.’

We proceed similarly for null subjects, which are also claimed to occur under a V2 rule, at least in the earliest period (for OFr see Roberts 1993, Adams 1987; for OFI see Benincà 1983/84, 2006, Poletto 2020b; for OVen see Poletto 2020a). In (3), the OVen sentence contains two main clauses with a null subject is followed by a

¹ In glossing the examples, we adhere to the Leipzig Glossing Rules. As most Venetian verbs are syncretic between 3SG and 3PL, we mark them as either 3SG or 3PL depending on the features of the subject.

coreferential overt subject (*el* ‘he’) in the next subordinate clause. In (4), we observe a main-clause null subject in French.

- (3) Old Venetian, Lio Mazor 18 (1311)
 ... et levà lo rem et de=me su lo col et
 ... and raised.3SG the oar and gave.3SG=DAT.1SG on the neck and
 menà=me ço per lo braço, sì ch’ el me lo
 hit.3SG= OBJ.1SG down for the arm so that he DAT.1SG OBJ.3SG
 scaveçà.
 broke.3SG
 ‘... and he raised his oar and hit me on the neck and stroke me down the arm
 so that he broke it.’
- (4) Old French, Atiremens (1314)
 En jugement se mistrent.
 in judgement SELF put.PST.3PL
 ‘They put themselves in judgement.’

Finally, the same reasoning is applied to the placement of clitic complement pronouns after the finite verbs in Venetian, following the so-called “Tobler-Mussafia law”, as this phenomenon has also been reported to be confined to main clauses in Venetian (Benincà 1983/84, 2006). Already (3) presents two cases of enclisis, while an additional occurrence is reported in (5).²

- (5) Old Venetian, Lio Mazor 2 (1311)
 ... e mis=me man en cavo.
 ... and put.PST=DAT.1SG hand on head
 ‘... and he put his hand on my head.’

If there is a correlation between these phenomena and V2, the expectation is that it will be detected in a similar evolution of rates of use and of asymmetrical distribution through time.

The data used to test this hypothesis are extracted from a treebank of two Medieval Romance varieties, French and Venetian, which share several grammatical evolutions, including the loss of V2. The treebank offers ample data points, since it comprises one text per century for the 14th, 15th, and 16th century. To enhance comparability, the treebank contains prose texts from the same region (Normandy for French and the Venetian Lagoon for Venetian) and belonging to the non-literary legal genre. The choice is driven by the fact that this genre has been found through preliminary investigations to yield less conservative rates of use of changing variables (Larrivé 2022; Goux and Pinzin in press).

² The same observation on clitic placement has been found to hold for other Old Romance varieties as Old Florentine (Benincà 1995), Old Spanish (Wanner 1991, Fontana 1993) and Old Catalan (Fischer 2002, 2003). Old French had the same phenomenon, but only up to the end of the 13th century, a period outside the scope of our contribution (Ramsden 1963).

a. <i>Lio Mazor</i>	1311-12	≅14000 tokens	Venetian
b. <i>Senato registri terra</i>	1440-1455	≅36000 tokens	Venetian
c. <i>Statuti di Murano</i>	1502	≅17000 tokens	Venetian
d. <i>Atiremens et Jugies</i>	1314	≅15000 tokens	French
e. <i>Style de l'Échiquier</i>	1425	≅71000 tokens	French
f. <i>Style Terrien</i>	1578	≅40000 tokens	French

The corpus is tokenized, lemmatized, with POS and syntactic function annotation (<https://txm-crisco.huma-num.fr/txm/>).³

The annotated texts are systematically considered for position of the finite verb in main and subordinate clauses, using parallel extraction queries capitalizing on the fine-grained PENN annotation set (Santorini 2021) which is sensitive to phrase boundaries and phrase-structure, further assisted by the graph-based dependency parser HOPS for the old French texts (Grobol & Crabbé, 2021).⁴

The extraction process, however, raises methodological questions related to how some configurations are counted. Coordinated clauses, for example, present a high share of V1 but mainly because of cases like (6)-(7), where the subject of the two conjoined clauses is the same.

(6) Old Venetian, Lio Mazor 1 (1311)

E en questa el esì fora de sot el portego e
 and in this he went.3SG out of under the portego and
 ven=me encontra.
 came.3SG=DAT.1SG towards
 ‘Meanwhile, he came out of the *portego* and came toward me’

³ Because of differential availability of tools, the French and Venetian are initially tackled by different POS/syntactic annotation systems, that are subsequently aligned by transformation scripts.

⁴ The extraction queries, the results of the queries and the final excel tables from which the graphs in Section 3 have been generated are available here: https://osf.io/s82gz/?view_only=2289379a74dc43dd9c6469b018e11462. Since the French texts had been annotated automatically (using BertForDeprel parser and a model trained on Profiterole corpus and finetuned with a corrected sample of 100 sentences from *Atiremens et Jugies*) and the corpus does not yet have a “Gold” status, to approach an adequate representation of the language and limit manual intervention, we extracted and then manually reviewed the annotation for the verb position and subject realisation/position. While we tagged and reviewed all main clauses, we decided in the interest of time to automatically select a representative sample of 500 dependent clauses per text using a Python script. These 500 clauses emerged from an initial automatic extraction of 300 clauses annotated as complete, 300 clauses annotated as adverbial, which were manually examined to eliminate inadequately annotated cases, and leaving aside any target attestations beyond the number of 500.

- (7) Old French, *Échiquier* (1425)
 ... leurs advocas comparent devant le juge et plaident ung
 ... their lawyers appear.3PL before the judge and plead.3PL a
 jugement.
 judgment
 ‘... their lawyers appear before the judge and plead a judgment.’

The same is true for subject and object relative clauses, which present a high share of V1 either because of a subject gap – as in subject relative clauses (8)-(9) – or because of the postverbal position of the subject – as with object relative clauses (10)-(11).

- (8) Old Venetian, *Lio Mazor 1* (1311)
 ... la palada ch' era averta de not.
 the tower that was.3SG open of night
 ‘... the tower that was open at night.’
- (9) Old French, *Atiremens* (1314)
 ... un homme qui achate terre.
 a man who buys land
 ‘... a man who buys land.’
- (10) Old Venetian, *Lio Mazor 23* (1311)
 ... che narò per saver que vol li cunsegeri.
 that go.FUT.1SG to know.INF that want.3PL the counsellors
 ‘... that I will go to get to know what the counsellors want.’
- (11) Old French, *Échiquier* (1425)
 ... le propos que demandoit son adversaire.
 the discourse that asked.3SG his adversary
 ‘... the project that his adversary asked.’

Both coordinated clauses and subordinates introduce a subject relative marker have therefore been set aside for the analysis. The results of the analysis of the (a)symmetry of V2 phenomena are presented in the next section.

3. Results

The purpose of this section is to report on the evolution of distributional clausal asymmetry of V2 and V2-related phenomena in Medieval French and Venetian from the beginning of the 14th to the 16th century. The overall progression of evolution in the strictly comparable corpora is reported, and commented upon in view of their relation to the hypothesis that the loss of V2 should be reflected in the loss of dominant main-clause distribution of other V2-dependent phenomena. The first phenomenon reported upon is the overall position of the verb through the texts.

3.1. Linear word order

The first measure of V2 change is the linear word-order patterns found across texts in each language. The rates of V1, V2, V3 and V4 is established via queries retrieving the relative position of the verb with respect to the other clausal constituents. Following standard assumptions, clitics (direct/indirect object, locative, impersonal, reflexives, negation) and coordinating particles (like *and/or/but*) are not counted as constituents, so that both (12) and (13) are considered as having the finite verb in second position in the clause (*possa* ‘can’ and *dura* ‘lasts’, respectively).⁵ Equally, elements identified as tokens in the treebank – punctuation, null subjects – that do not contribute overt structural material were not counted.

- (12) Old Venetian, Senato Terra Registri f.29v (27 May 1446)

Item non se possa i dicti dacia de sal
 equally not SELF can.SBJV.3SG the said duties of salt
 inchantar. (OVen)
 auction.INF
 ‘Equally, the said salt duties cannot be auctioned.’

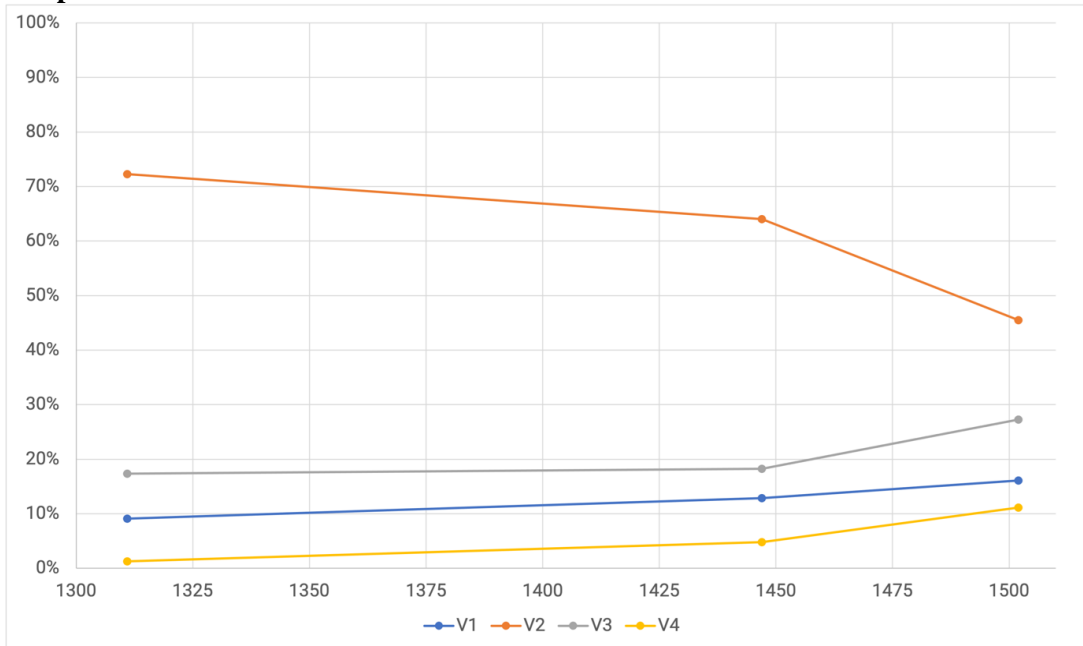
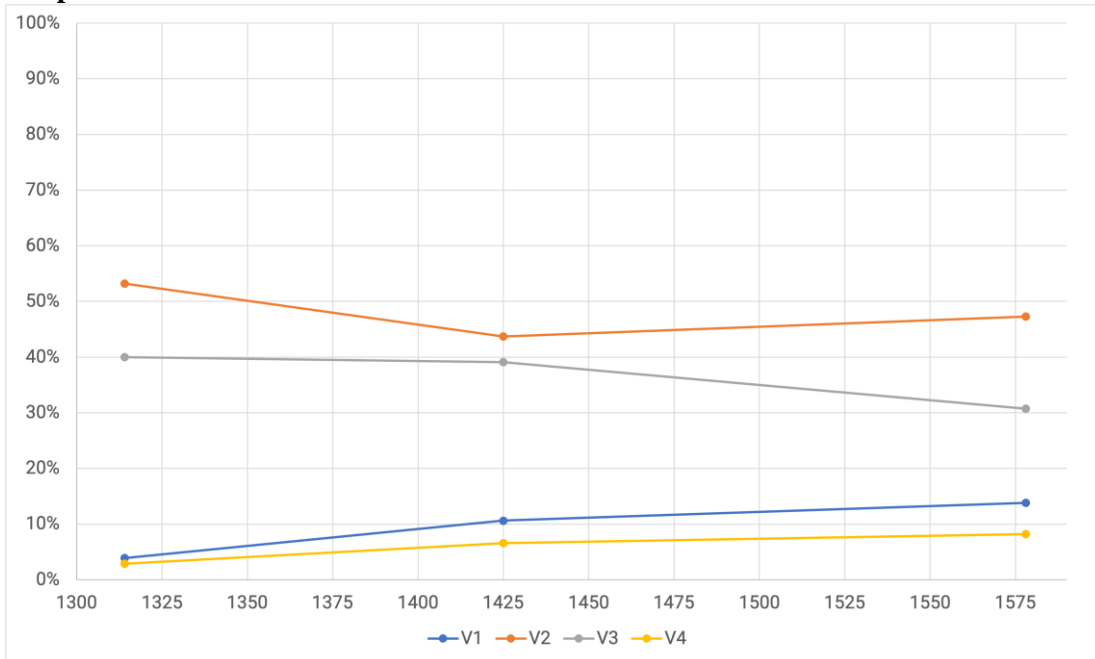
- (13) Old Venetian, Senato Terra Registri f.66v (14 June 1453)

E questo dura per fin le dicte donzelle se maridera
 and this lasts for end the said maidens SELF marry.FUT.3PL
 ‘and is valid until the maidens mentioned above will get married [...]’

Overall, Venetian is increasingly more V3/V4 (from 18,6% to 23,1% to 38,4%) and less V2 (from 72,3 to 64% to 45,5%). This is presented in Graph 1.⁶ On the contrary, the share of V2 is stable in French (from 53,2% to 43,7% to 47,3%), V1 and V4 increase (from 3,9% to 10,6% to 13,8% and from 2,7% to 5,8% to 7,5%, respectively) and V3 decreases (from 40% to 39,1% to 30,7%). This is shown in Graph 2.

⁵ We did not consider in the French sample clauses introduced by *car*, as its analysis is controversial; it is considered as a coordinator by grammatical tradition, despite its behaviour as a subordinator.

⁶ V3 and V4 are not merged in a single category because they frequently display different paths. This is the case, for example, in Old High German, where V3 is consistently attested – with specific initial XP constituents – while V4 (or more) is almost absent (see Fuß 2008 and references therein for an overview).

Graph 1. Venetian Word Order**Graph 2. French Word Order**

Comparing the two languages, one can see how the Venetian proportions diachronically converge toward the French ones, which are rather diachronically stable. In the earliest Venetian text, V2 is by far the most widespread configuration (around 70%). This changes through time, converging toward the French situation, which sees main clauses presenting linear V2 around 50% of the time already from the beginning of the 14th century. Although quantitative assessment of linear V2 is not a guarantee that the language in question is governed by a grammar involving movement of the finite verb to the C layers (V-to-C) – as SVO too can give rise to linear V2 – getting away from strict V2 order for languages known to display the V2

phenomenology in the initial period is a good indication that the V2 rule is being lost. This is indeed the case for Venetian, while French does not show any clear diachronic evolution.

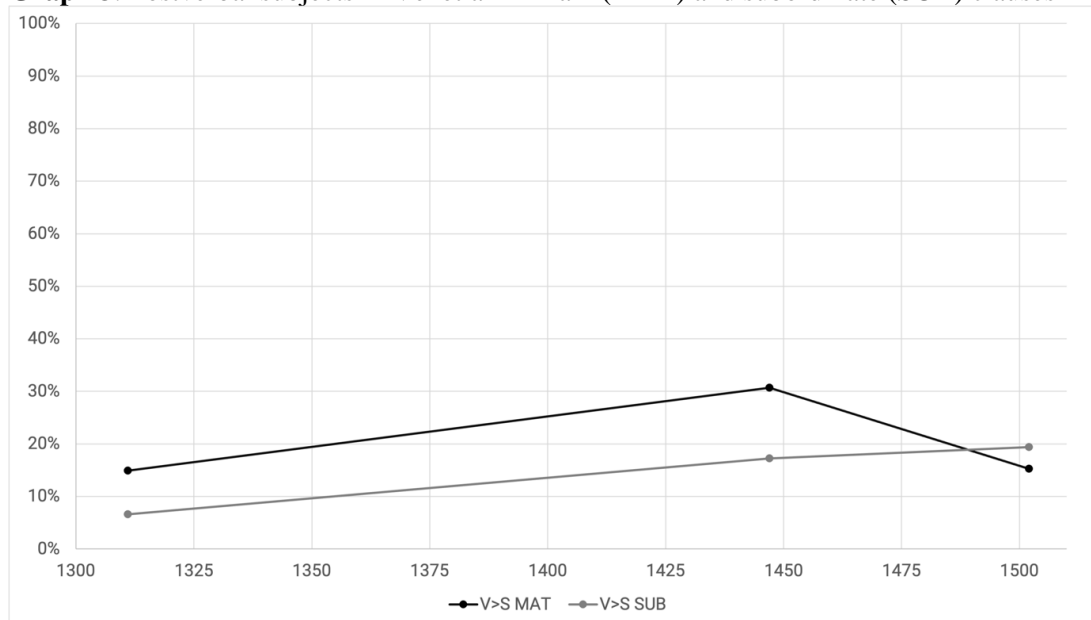
This measure serves as a starting point to check whether there is a correlation with the behavior of other phenomena thought to relate to V2. The expectation would be that asymmetry of these phenomena would be lost in Venetian, which as we just saw loses predominant V2 order, but not in French, which is largely stable in this respect. The phenomenon that we report upon next is subject inversion.

3.2. Subject inversion

As stated in the introduction, the question we investigate in this work is whether the distributional asymmetry of V2 is correlated to that of other related phenomena. One phenomenon generally thought to depend on V2 word order is subject inversion. As V2 is a main-clause phenomenon, it should mostly distribute in main clauses, if indeed subject inversion is linked to the V2 rule. This dominance should be lost as V2 word order is. In light of the previous Section, we therefore expect Venetian to show a diachronic reduction of the main/subordinate asymmetry, while we expect stability for French.

Subjects in postverbal position were extracted in each language from main and subordinate clauses – exclusive of coordinated main clauses and subordinate relative clauses – and relativized to the total of subjects (preverbal, postverbal or null) in the relevant clause-type. The results for Venetian are in Graph 3.

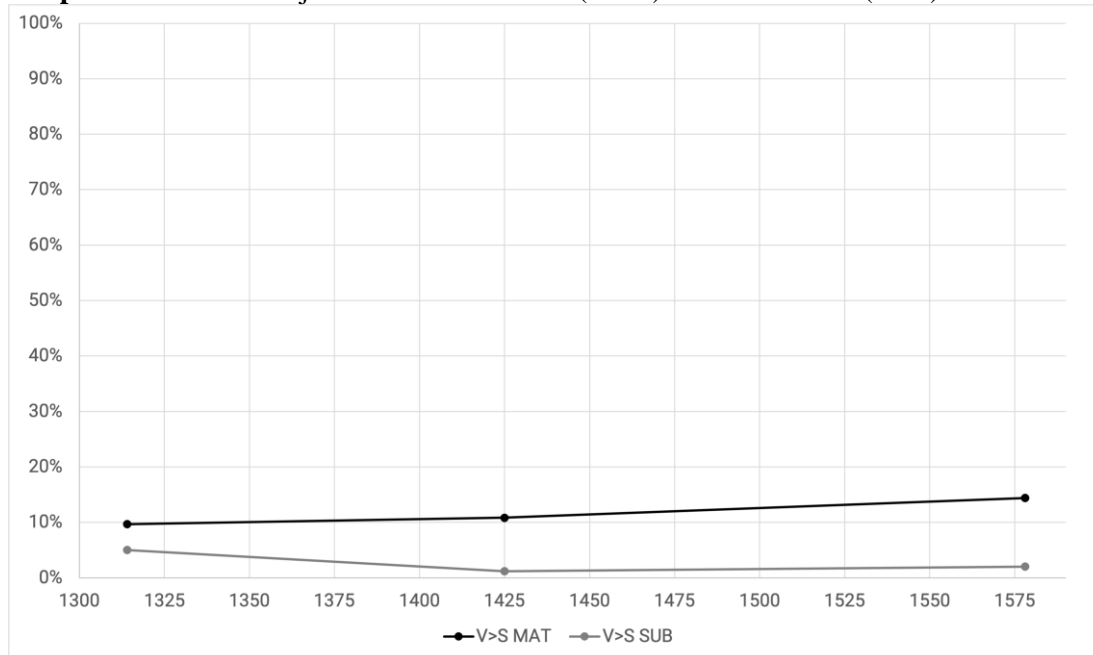
Graph 3. Postverbal subjects in Venetian in main (MAT) and subordinate (SUB) clauses



Graph 3 presents the proportion of relevant main clauses and of subordinate clauses with postverbal subjects in our Venetian texts (*Lio Mazor*: 14,9% main vs. 6,6% subordinate; *Senato terra*: 30,7% main vs. 17,2% subordinate; *Murano custom*: 15,3% main vs. 19,4% subordinate). Up to the middle of the 15th century, postverbal subjects increase through time and maintain a dominant distribution in main clauses.

The difference between rates of subject inversion in main and subordinates increases from 8,3% to 13,5%. Due to the rising rates, however, the ratio is slightly decreasing ($14,9/6,6 = 2,3$ vs. $30,7/17,2 = 1,8$). In contrast with the first two texts, the last text shows a reversed pattern, with an advantage of subordinate clauses over main clauses with respect to subject inversion ($15,3/19,4 = 0,8$). Overall, we see that – after an initial reduction of the advantage of postverbal subjects in main clauses ($2,3 > 1,8$) – postverbal subjects completely lose their asymmetrical distribution in Venetian. Whether this is the case in French is made visible by Graph 4.

Graph 4. Postverbal subjects in French in main (MAT) and subordinate (SUB) clauses



Graph 4 presents the proportion of relevant main clauses and of subordinate clauses with postverbal subjects in our French texts (*Atirements*: 9,7% main vs. 5% subordinate; *Style Echequier*: 10,8% main vs. 1,2% subordinate; *Style Terrien*: 14,4% main vs. 2,1% subordinate). Contrary to the Venetian texts, the French texts show an increase in the asymmetrical distribution of postverbal subjects between the first and the last two texts, from **1,9** (9,7/5) to **9** (10,8/1,2) / **6,9** (14,4/2,1). All in all, we see that the asymmetry is maintained through time, even with increasing numbers. No clear diachronic path is visible.

Comparing the two datasets, we notice that the overall frequency of postverbal subjects is higher in Venetian than in French. This could indicate a difference in terms of the available sources for postverbal subjects. It is well known that postverbal subjects can have different sources, not only inversion due to V-to-C movement (V2), but also the availability of vP peripheral focus (Belletti 2004) and of a low VP-internal position (for subjects of unaccusatives, passives, etc. see Belletti 1988, Burzio 1986, Pinto 1994 among others). Given the higher rate of postverbal subjects in Venetian both in main and in subordinate clauses, we interpret this higher frequency of postverbal subjects as depending on a new possibility of inversion that is not related to V2. We surmise that the data of the last Venetian text must be interpreted as a consequence of the rising of vP peripheral subjects, which are not possible in the

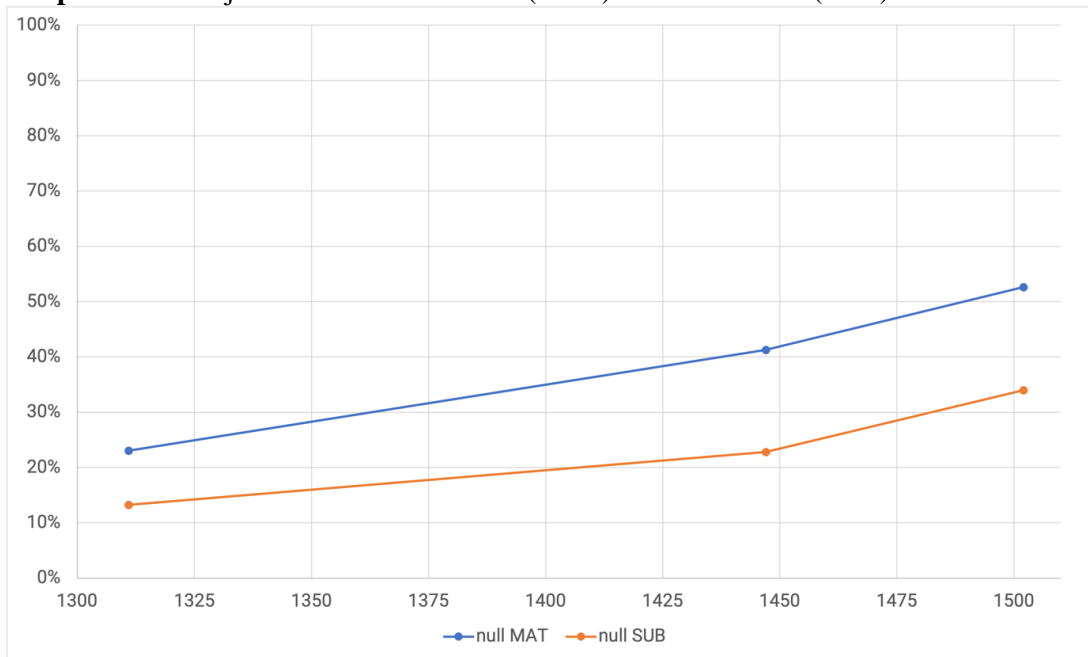
earlier periods but are possible in the modern variety.⁷ The same cannot be said of the French texts, as postverbal subjects in subordinate clauses follow a rather flat line, occurring at very low rates in each of the three texts investigated.

Notwithstanding this, the relevance of the main-subordinate asymmetry for the distribution of postverbal subjects is visible in both languages, except for the last Venetian text (1502). In other words, we see that in Venetian postverbal subjects stop being asymmetrically distributed only in the last text (1502), while the asymmetrical distribution is always maintained in French. This indicates that V2 plays a significant role in the distribution of postverbal subjects for the whole period in French and up until the middle of the 15th century for Venetian. The main/subordinate asymmetry gets therefore lost in Venetian and maintained in French. In this sense, this supports the view that the V2 word-order rule is losing ground in Venetian, as shown in the previous section. French, on the other hand, does not seem to undergo the same diachronic trajectory, also this in line with the observations in the previous section. The next section looks at the case of null subject.

3.3. Null subjects

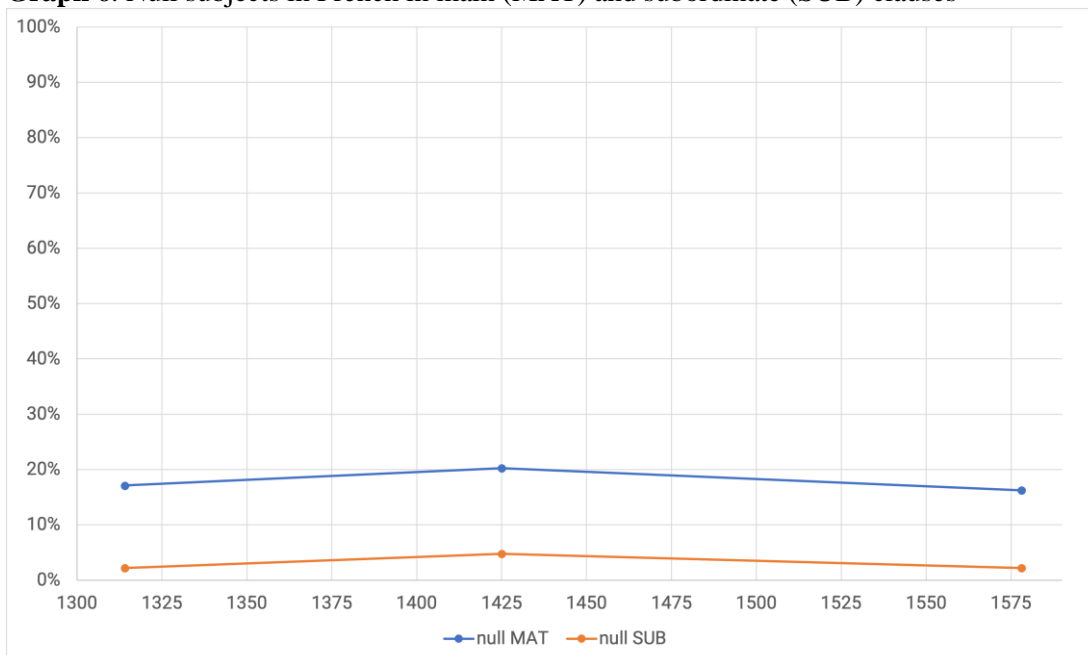
As pointed out in Section 2, it is widely believed that null subjects in the investigated languages are at least initially dependent on a V2 rule. As such, a distributional asymmetry should be found between main and subordinate clauses in the two languages. To check for this, we extracted the cases of null subjects in the two languages from the total of main and subordinate clauses, exclusive of coordinated main clauses and subordinate relative clauses. Let us look first at the figures from Venetian, as per Graph 5.

⁷ Note that if we compare the start and end point, the overall situation of subject inversion in subordinate clauses in Venetian and French are opposed in to their progression, rising in the second and declining in the latter. This is in line with the fact that the Modern varieties differ in the availability of both postverbal focused subjects in the vP area and postverbal subjects in the VP, with French allowing only for VP-internal indefinite subjects (Belletti 2004).

Graph 5. Null subjects in Venetian in main (MAT) and subordinate (SUB) clauses

Graph 5 presents the rate of null subjects on the total of relevant main and subordinate clauses for Venetian (*Lio Mazor*: 23% main vs. 13,2% subordinate; *Senato terra*: 41,2% main vs. 22,8% subordinate; *Murano custom*: 52,6% main vs. 34% subordinate). More strikingly than with subject inversion, there is a rise of null subjects in Venetian, both in main and subordinate clauses. Independently of this observation, the ratio of the share of null subjects in main vs. subordinate clauses is stable between the first and the second text ($23/13,2 = 1,74$; $41,2/22,8 = 1,81$), while it slightly decreases for the third ($52,6/34 = 1,55$).

We now look at the overall French situation and the summarizing Graph 6.

Graph 6. Null subjects in French in main (MAT) and subordinate (SUB) clauses

Graph 6 presents the proportion of relevant main clauses and of subordinate clauses with null subjects in our French texts (*Atirements*: 17,1% main vs. 2,2% subordinate; *Style Echequier*: 20,3% main vs. 4,8% subordinate; *Style Terrien*: 16,2% main vs. 2,2% subordinate). Again, the French evolution appears distinct from that of Venetian. Instead of an increase, we see a stable occurrence of null subjects, always asymmetrically distributed between main and subordinate clauses. Comparing the first text with the last two, the ratio of null subjects in main vs. subordinate clauses is perfectly stable ($17,1/2,2 = 7,7$ vs. $20,3/4,8 = 4,2$ and $16,2/2,2 = 7,4$). This is in line with the French data on both verb position and postverbal subjects (Section 3.1 and 3.2), with no manifest diachronic evolution.

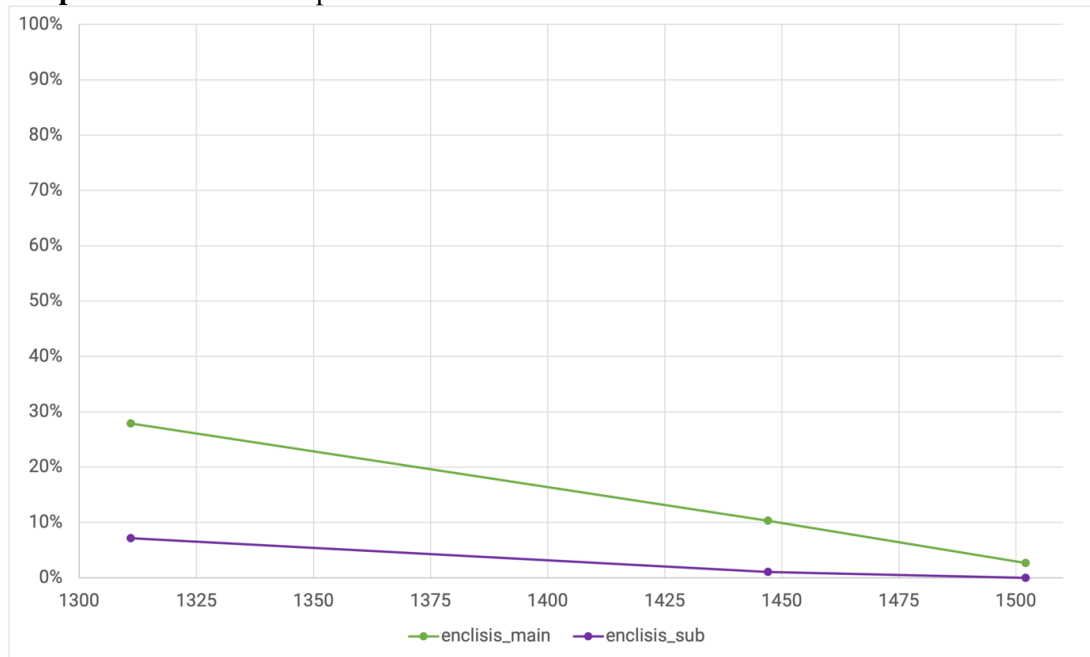
The evolution of null subjects in Venetian and French thus diverges. In French null subjects are a stable main clause phenomenon, while in Venetian they increase overall, both in main and subordinate clauses.⁸ Nonetheless, while Venetian shows a gradual loss of the asymmetrical distribution of null subjects between main and subordinate clauses, French shows diachronic stability. This again supports the idea that loss of asymmetry, as a cue to V2, reflects loss of V2, and is refracted in other phenomena.

Another peculiarity of Venetian (and other Old Romance varieties, see footnote 1) is explored in the next section.

3.4. Venetian enclisis

An almost exclusively main clause phenomenon found to correlate with V2 is enclisis to the finite verb, by which the finite verb precedes its complement clitics in a declarative context when there is V1 (see Section 2). Enclisis to main declarative verbs is a well-attested phenomenon across most Medieval Romance varieties, it is however singularly absent from French, at least from the end of the XIII c. (see footnote 1). Indeed, the French corpus contains no occurrence of it. Considering this, we focus here only on the Venetian data, presented in Graph 7.

⁸ A plausible explanation for the rise in the share of null subjects in Venetian might be the influence of Tuscan, where null subjects are more widely attested than in French already by the 13th century. (see the discussion in Poletto 2020b). Further support for the Tuscan influence is the frequent presence in the later texts of an overt morphological difference between 3SG and 3PL verbal forms. This feature characterizes Tuscan and not Venetian varieties, which show a generalized syncretism (but for specific strong forms).

Graph 7. Rate of enclisis per main ad subordinate clauses with a clitic in Venetian.

Graph 7 makes visible the share of enclisis to the finite verb on the total of occurrences in which we find a clitic pronoun and a finite verb in the three venetian texts analyzed. As the graph clearly shows, enclisis in main clauses sharply drops after the initial period. It goes from 27,9% in the *Lio Mazor* text, to 10,3% in the *Senato terra* text, to 2,7% in the *Murano* text. Even more relevantly, it seems that it loses productivity, occurring in the later texts almost exclusively with the subjunctive form of the verb *dover* ‘must’ in combination with the clitic *se* ‘self’ (14), which could be interpreted as a fixed form.

- (14) Old Venetian, Murano I, 2, a (1502)
 ... et debia=se far la election di predicti XXX in questo
 ... and must=SELF do.INF the election of.the before.said thirty in this
 modo.
 manner
 ‘... and the election of the already mentioned XXX must proceed in this way.’

In subordinate clauses the rate is always extremely low (6 occurrences = 7,1%; 2 occurrences = 1,1%; 0 occurrences = 0%) and, more importantly, the 8 occurrences we find in the first two texts are all present in a coordinated subordinate clause with no overt repetition of the complementizer (15).

- (15) Old Venetian, Lio Mazor 4 (1311)
 ... e videsem che Nicolò çunse Piçol Pare e
 ... and saw.1PL that Nicolò reached.3SG Piçol Pare and
 començà=li menar d' algudola e de'=li;
 began.3SG=DAT.3SG hit.INF of knife and gave.3SG=DAT.3SG
 ‘... and we saw that Nicolò caught up with Piçol Pare and started hitting him
 with the knife and beat him up;’

Setting aside these 8 problematic occurrences, the asymmetry is categorical: enclisis to the finite verb is only attested in main clauses. Notwithstanding this last point, the relevant observation for this phenomenon is its general demise. It basically disappears in the later texts, except for some fixed forms. The main vs. subordinate asymmetry is categorical in the last text, while for the first two texts it rises (27,9/7,1 = **3,9**; 10,3/1,1 = **9,4**). This contrasts with what we observed for postverbal and null subjects in Venetian, where the asymmetry grows smaller through time while the overall frequency of the phenomenon increases. In this sense, the overall frequency of Venetian enclisis directly covaries with the loss of V2. As such, this phenomenon can be used as direct predictor for the (loss of) V2 phenomenology.

4. Conclusion

This work has sought to articulate a developmental scenario for the loss of V2, based on the systematic investigation of calibrated data documenting two Medieval Romance varieties from the 14th century to the 16th century. V2 is measured by the proportion of main and embedded clauses where the verb is in second position. In Venetian, the rate of V2 word order decreases, from 72% to 46% (with a correlative rise of V1 and V3+ from 19% to 38%); whereas in French it remains stable at around 50%. In view of this, because the V2 rule is a main clause phenomenon, we expect the V2-related phenomena of subject inversion, null subject and enclisis – assessed by comparing numbers in main clauses and subordinates through time – to lose their asymmetrical distribution in Venetian, and not in French. This is indeed the case, as shown in Table 1 for Venetian and in Table 2 for French.

Table 1. Share of Verb Second and V2-related phenomena, Venetian

text	year	V2	Subject inversion		Null subjects		enclisis	
			main	sub	main	sub	main	sub
<i>Lio Mazor</i>	1311-1312	72,3%	14,9%	6,6%	23%	13,2%	27,9%	7,1%
<i>Senato terra</i>	1440-1455	64%	30,7%	17,2%	41,2%	22,8%	10,3%	1,1%
<i>Statuti Murano</i>	1502	45,5%	15,3%	19,4%	52,6%	34%	2,7%	0%

Table 2. Share of Verb Second and V2-related phenomena, French

text	year	V2	Subject inversion		Null subjects	
			main	sub	main	sub
<i>Atiremens et Jugies</i>	1314	53,2%	9,7%	5%	17,1%	2,2%
<i>Style de l'Échiquier</i>	1425	43,7%	10,8%	1,2%	20,3%	4,8%
<i>Style Terrien</i>	1578	47,3%	14,4%	2,1%	16,2%	2,2%

A progressive increase in subject inversion in embedded clauses is found in Venetian, with an initial ratio of 2,3 times of main clause to subordinate clause subject inversion, to a final inverted ratio of 0.8, where subordinate clauses show more subject inversion than main clauses. In line with the expectations, the proportion of subject inversion in main vs. subordinate clauses is mostly stable in French, with a constant advantage of main clauses (from 1,9 times to 9 times, to 6,9 times). For null subjects, the ratio of main to subordinate goes in Venetian from an initial 1,74 to a final 1,55, showing the expected reduction of the advantage of main clauses (potentially mitigated by an overall orthogonal increase in null subjects). In this respect, French is coherently diachronically consistent, in the sense that it shows a constant advantage of main clauses for all our three texts (from 7,7 times to 4,2 times, to 7,4 times). For Venetian enclisis, finally, the ratio in main vs. subordinate clauses is increasing in between the first two texts (from 3,9 to 9,4) and not assessable in the last text due to lack of occurrences in subordinates. On the other hand, the overall frequency of the phenomenon is sharply decreasing, as expected if the overall phenomenon depends on the V2 property.

The detailed quantitative analysis of calibrated data suggests that linear V2 word order in main clauses is gradually replaced by V3+ through the Medieval period in Venetian, while French shows a constant rate with no clear diachronic evolution. On the asymmetry of related phenomena, the tendency is a reduction of the ratio of main to subordinate as the rate of linear V2 in main clauses decreases, as shown by the Venetian data. In keeping with the analysis, the French data show diachronic consistency in the asymmetrical occurrence of the related phenomena, always more frequent in main than in subordinate clauses with no reduction of such an asymmetrical distribution.

Finally, we would like to point out that even if all V2-related phenomena vary through time in Venetian, as expected looking at the share of linear V2 in main clauses, their diachronic paths are not the same. We observe, at least, two different types of curves. On the one hand, Venetian enclisis show a decrease in the global frequency of the phenomenon and an increase in the ratio of main to subordinate distribution (up to where the ratio can be calculated). As it involves small (or nil) numbers of occurrences, more data might therefore be useful in confirming the tendencies identified here. On the other hand, Venetian subject inversion and null subjects show a surge in the global frequency of the phenomenon and a decrease in the ratio of main to subordinate distribution. For both phenomena, this might be due to the diversification of their grammatical motivations.

Nonetheless, the significance of the investigation is that the move from V2 to SVO is diagnosed not only by the rates of main-clause subject inversion, but also by the rates of pro-drop asymmetry between main and subordinates and the loss of enclisis

to the inflected verb. This is convergent with an overall narrative by which high verb movement is lost gradually rather than in one go, undergoing a phase of movement to a low C position, rather than a brutal change to SVO. We expect such a narrative to be supported by the analysis of the evolution of other cues of V2, highlighting the necessity of a multifactorial analysis by which V2 is acquired and lost through a variety of cues.

Acknowledgments

This research is part of the “Micro Cues for Language Evolution” project (MICLE), jointly funded by the *Deutsche Forschungsgemeinschaft* and the *Agence Nationale de la Recherche* (ANR-20-FRAL-0001-01 / DFG-449439301). We thank these organizations for their financial support. The contribution greatly benefited from conversations with Sam Wolfe, Marc Olivier, Espen Klævik-Pettersen and the audience of the Linguistic Symposium on Romance Languages 2023. The initial scripts for tagging and parsing the Old Venetian data have been provided by Emanuela Sanfelici, whose help has been crucial in many respects. We would also like to thank two anonymous reviewers for their constructive comments and criticism. All errors remain our own.

Author’s contribution

Francesco Pinzin is responsible for the data collection and tagging of the Venetian data, Mathieu Goux and Pierre Larrivé are responsible for the data collection and tagging of the French data. The analysis has been developed by Pierre Larrivé, Cecilia Poletto and Francesco Pinzin. Pierre Larrivé and Francesco Pinzin are responsible for the first draft, which has then been reviewed by Cecilia Poletto.

Direct sources

Lio Mazar = Elsheikh, Mahmoud S. (ed.). 1999. *Atti del podestà di Lio Mazar, edizione critica*. Venezia: Istituto Veneto di scienze, lettere ed arti.

Senato registri terra = Archivio di Stato di Venezia [= ASVe], Senato, Deliberazioni, Terra, Registri, 1-3.

Statuti di Murano = Ortalli, Gherardo, Pasqualetto, Monica, Rizzi, Alessandra (eds.). 1989. *Statuti della Laguna Veneta dei secoli XIV-XVI: Mazzorbo (1316), Malamocco (1351-1360), Torcello (1462-1465), Murano (1502)*. Roma: Juvence.

Atiremens et Jugies = *Atiremens et Jugiés d'Eschiquiers*, transcribed by Robert Génestal & Ernest-Joseph Tardif, 1921, Caen : Tome IV de la Bibliothèque d'histoire du droit normand. Google Books ID: -vk7AQAAMAAJ

Style de l'Échiquier = “Coustume, stille et usage au temps des echiquiers de Normandie”. *Style et usage de l'échiquier de Normandie*, 1851, Paris : Société des Antiquaires de Normandie, p. 345-413. *Gallica*: [ark:/12148/bpt6k200065b](https://gallica.bnf.fr/ark:/12148/bpt6k200065b)

Style Terrien = *Commentaires du droict civil tant public que privé observé au pays et duché de Normandie - Style de procédure*, Guillaume Terrien, 1578, p.339-402. Paris: J. Du Puys. *Gallica*: [ark:/12148/bpt6k9107304x](https://gallica.bnf.fr/ark:/12148/bpt6k9107304x)

References

BertForDeprel = Guiller, Killian. 2020. Analyse syntaxique automatique du pidgin-créole du Nigeria à l'aide d'un transformer (BERT): Méthodes et Résultats. Paris 3 Sorbonne Nouvelle.

Profiterole = Prevost Sophie, Grobol Loïc, Dehouck Mathieu, Lavrentiev Alexeï, & Heiden Serge. 2024. Profiterole: Un corpus morpho-syntaxique et syntaxique de français médiéval *Corpus* (25) 16 pp.

Adams, Marianne. 1987. From Old French to the Theory of pro-drop. *Natural Language and Linguistic Theory* 5(1): 1–32. <https://doi.org/10.1007/BF00161866>

Belletti, Adriana. 1988. The Case of Unaccusatives. *Linguistic Inquiry* 19(1): 1–34. <http://www.jstor.org/stable/4178572>

Belletti, Adriana. 2004. Aspects of the Low IP Area. In L. Rizzi (ed.), *The Structure of CP and IP: The Cartography of Syntactic Structures Volume 2*, 16–51. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780195159486.003.0002>

Benincà, Paola. 1983/1984. Un'ipotesi sulla sintassi delle lingue romanze medievali. *Quaderni Patavini di Linguistica* 4: 3–19.

Benincà, Paola. 1995. Complement Clitics in Medieval Romance: The Tobler-Mussafia Law. In A. Battye, & I. Roberts (eds), *Clause Structure and Language Change*, 325–344. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780195086324.003.0012>

Benincà, Paola. 2006. A detailed map of the left periphery of Medieval Romance. In R. Zanuttini, H. Campos, E. Herbunger, & P. H. Portner (eds), *Crosslinguistic research in syntax and semantics. Negation, tense and clausal architecture*, 53–86. Washington: Georgetown University Press. <https://www.jstor.org/stable/j.ctt2tt7c3.6>

Benincà, Paola. 2013. Caratteristiche del V2 Romanzo. Lingue Romanze Antiche, Ladino Dolomitico e Portoghese. In E. Bidese, & F. Cognola (eds), *Introduzione alla linguistica del mòcheno*, 65–84. Turin: Rosenberg & Sellier.

Burzio, Luigi. 1986. *Italian syntax*. Dordrecht: Reidel. <https://doi.org/10.1007/978-94-009-4522-7>

Cichosz, Anna, Gaszewski, Jerzy, & Piotr Pezik. 2016. *Element Order in Old English and Old High German Translations*. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/nss.28>

Diesing, Molly. 1990. Verb movement and the subject position in Yiddish. *Natural Language and Linguistic Theory* 8(1): 41–79. <https://doi.org/10.1007/BF00205531>

Fischer, Susann. 2002. *The Catalan Clitic System: A diachronic perspective on its syntax and phonology*. Berlin & New York: Mouton de Gruyter. <https://doi.org/10.1515/9783110892505>

Fischer, Susann. 2003. Rethinking the Tobler-Mussafia Law. Data from Old Catalan. *Diachronica* 20(2): 259–288. <https://doi.org/10.1075/dia.20.2.03fis>

Fontana, Josep M. 1993. *Phrase Structure and the Syntax of Clitics in the History of Spanish*. Ph.D. thesis, University of Pennsylvania.

Foulet, Lucien. (1929). *Petite syntaxe de l'ancien français*. Paris: Champion.

Fuß, Eric. 2008. *Word order and language change. On the interface between syntax and morphology*. Habilitationsschrift, Goethe-Universität Frankfurt.

Goux, Mathieu, & Francesco Pinzin. In press. Challenges of a Multilingual corpus (Old French/Old Venetian): The example of the *MICLE* project. In E. Castro, A. Della Fontana, & E. Pezzini (eds), *Venise et la France. Similitudes, spécificités, interrelations*. Firenze: Franco Cesati Editore.

Grobol, Loïc, & Benoît Crabbé. 2021. Analyse en dépendances du français des plongements contextualisés. In *TALN-RÉCITAL 2021*. https://hal.archives-ouvertes.fr/hal-03223424/file/HOPS_final.pdf.

Holmberg, Anders. 2015. Verb Second. In A. Alexiadou (ed.), *Syntax - Theory and Analysis, Volume I*, 342–382. Berlin/München/Boston: De Gruyter Mouton. <https://doi.org/10.1515/9783110377408.342>

Kaiser, Georg A. 2002. *Verbstellung und Verbstellungswandel in den romanischen Sprachen*. Tübingen: Niemeyer. <https://doi.org/10.1515/9783110911640>

Kaiser, Georg A., & Michael Zimmermann. 2011. On the decrease in subject-verb inversion in French declaratives. In E. Rinke, & T. Kupisch (eds), *The development of grammar: Language acquisition and diachronic change: In honour of Jürgen M. Meisel*, 355–381. Amsterdam/ Philadelphia: John Benjamins. <https://doi.org/10.1075/hsm.11.19kai>

-
- Labelle, Marie. 2007. Clausal architecture in Early Old French. *Lingua* 117(1): 289–316. <https://doi.org/10.1016/j.lingua.2006.01.004>
- Larrivé, Pierre. 2019. To be or not to be informational: Preverbal complements in Medieval French V2 configurations. *Glossa: a journal of general linguistics* 4(1): 1–19. <https://doi.org/10.5334/gjgl.851>
- Larrivé, Pierre. 2022. Is Medieval French diglossic? New evidence on remnant V2 and register. *Isogloss* 8(2): 1–16. <https://doi.org/10.5565/rev/isogloss.114>
- Le Coultre, Jean-Jules. 1875. De l'ordre des mots dans Crestien de Troyes. In E. Ziel (ed.), *XIV. Programm des Vitzhumschen Gymnasiums, als Einladung zu dem 17. 18 und 19. März 1875 stattfindenden öffentlichen Examen und Redeactus*, 5–88. Dresden: Teubner.
- Liu, Haitao. 2010. Dependency direction as a means of word-order typology: A method based on dependency treebanks. *Lingua* 120(6): 1567–1578. <https://doi.org/10.1016/j.lingua.2009.10.001>
- Mathieu, Éric. 2013. The Left-Periphery in Old French. In D. L. Arteaga (ed.), *Research on Old French: The State of the art*, 327–350. Dordrecht: Springer. https://doi.org/10.1007/978-94-007-4768-5_17
- Meisel, Jürgen M., Elsig, Martin, & Esther Rinke. 2013. *Language acquisition and change: A morphosyntactic perspective*. Edinburgh: Edinburgh University Press. <http://www.jstor.org/stable/10.3366/j.ctt9qdnj>
- Pinto, Manuela. 1994. Subjects in Italian: distribution and interpretation. In R. Bok-Bennema, & C. Cremers (eds), *Linguistics in the Netherlands 1994*, 175–186. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/avt.11.18pin>
- Poletto, Cecilia. 2014. *Word order in Old Italian*. Oxford: Oxford University Press.
- Poletto, Cecilia. 2019. More than one way out: On the factors influencing the loss of V to C movement. *Linguistic Variation* 19(1): 47–81. <https://doi.org/10.1075/lv.16001.pol>
- Poletto, Cecilia. 2020a. On the licensing of null subjects in Old Venetian. In H. Lohnstein, & A. Tsiknakis (eds), *Verb Second: Grammar Internal and Grammar External Interfaces*, 277–294. Berlin/Boston: De Gruyter Mouton. <https://doi.org/10.1515/9781501508141-009>
- Poletto, Cecilia. 2020b. Null subjects in Old Italian. In R. Woods, & S. Wolfe (eds), *Rethinking verb second*, 325–347. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198844303.003.0014>
- Ramsden, Herbert. 1963. *Weak-Pronoun Position in the Early Romance Languages* Manchester: Manchester University Press, Manchester.

Rinke, Esther. 2007. *Syntaktische Variation aus synchronischer und diachronischer Perspektive. Die Entwicklung der Wortstellung im Portugiesischen*. Frankfurt am Main, Madrid: Vervuert Verlagsgesellschaft. <https://doi.org/10.31819/9783964561930>

Rinke, Esther, & Jürgen Meisel. 2009. Subject inversion in Old French: Syntax and information structure. In G. A. Kaiser, & E.-M. Remberger (eds), *Proceedings of the Workshop Null-subjects, expletives, and locatives in Romance*, 93–130. Konstanz: Fachbereich Sprachwissenschaft, Universität Konstanz.

Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. In L. Haegemann (ed.), *Elements of Grammar. Handbook of Generative Syntax*, 281–337. Dordrecht/Boston/London: Kluwer academic publishers. https://doi.org/10.1007/978-94-011-5420-8_7

Roberts, Ian. 1993. *Verbs and Diachronic Syntax: A Comparative History of English and French*. Dordrecht/Boston/London: Kluwer academic publishers. <https://doi.org/10.1007/978-94-011-2910-7>

Rögnvaldsson, Eiríkur, & Höskuldur Thráinsson. 1990. On Icelandic word order once more. In J. Maling, & A. Zaenen (eds), *Modern Icelandic Syntax*, 1–40. Leiden: Brill. https://doi.org/10.1163/9789004373235_002

Santorini, Beatrice. 1992. Variation and change in Yiddish subordinate clause word order. *Natural Language & Linguistic Theory* 10(4): 595–640. <https://doi.org/10.1007/BF00133331>

Santorini, Beatrice. 2021. *Syntactic annotation manual for historical French*. <https://www.ling.upenn.edu/~beatrice/annotation-french>

Singh, Onkar. 2021. Verb Second in Old Venetian. *Isogloss* 7: 1–28. <https://doi.org/10.5565/rev/isogloss.110>

Sitaridou, Ioanna. 2019. Against V2 in Old Spanish. In A. Breitbarth, M. Bouzouita, L. Danckaert, & M. Farasyn (eds), *The determinants of diachronic stability*, 131–156. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.254.06sit>

Thurneysen, Rudolf. 1892. Zur Stellung des Verbums im Altfranzösischen. *Zeitschrift für romanische Philologie* 16(1/4): 289–307. <https://doi.org/10.1515/zrph.1892.16.1-4.289>

Vance, Barbara. 1997. *Syntactic change in medieval French: verb-second and null subjects*. Dordrecht: Kluwer academic publishers. <https://doi.org/10.1007/978-94-015-8843-0>

Walkden, George, & Hannah Booth. 2020. Reassessing the historical evidence of embedded V2. In R. Woods, & S. Wolfe (eds), *Rethinking verb second*, 536–554.

Oxford: Oxford University Press.
<https://doi.org/10.1093/oso/9780198844303.003.0022>

Wanner, Dieter, 1991. The Tobler-Mussafia Law in Old Spanish. In H. Campos, & F. Martínez-Gil (eds), *Current Studies in Spanish Linguistics*, 313–378. Washington, DC: Georgetown University Press.

Wolfe, Sam. 2018. *Verb Second in Medieval Romance*. Oxford: Oxford University Press.

Wolfe, Sam. In press. A conspiracy theory for the loss of v2 in Romance. *Journal of Historical Syntax*.

Woods, Rebecca, & Sam Wolfe (eds). 2020. *Rethinking Verb Second*. Oxford: Oxford University Press.