

Are French and Venetian V2 languages? A diachronic treebank analysis

Cecilia Poletto^{1,2}, Pierre Larrivé³, Francesco Pinzin¹, Mathieu Goux³

1: Goethe-Universität Frankfurt am Main; 2: Università degli Studi di Padova; 3: Université de Caen Normandie

It is widely assumed that the typologically rare V2 word order was operative in at least some of the Medieval Romance languages. This has been challenged however, for Ibero-Romance notably (Sitaridou et al. 2019), but also for French itself (especially in the work of Kaiser, e.g., 2002). Even looking at the most influential recent monographs (Meisel, Elsig & Rinke 2013, Poletto 2014, Wolfe 2018), the putative V2 status of Medieval Ibero-Romance, Gallo-Romance and Italo-Romance has tended to be assessed on the basis of representative examples from verse and prose literary texts from various regions. When systematic studies of a coherent empirical set have been performed, they have generally relied on extended passage of sources rather than sources as a whole for reasons of expediency. Furthermore, the main tests used in the literature to assess the V2 status have been the rate of subject inversion vs. subject initial clauses and the V2 linear restriction itself. This even though they have been shown to be problematic in view of the Romance configuration of the CP (Rizzi 1997) and the existence of vP peripheral subjects (Belletti 2004). Other phenomena related to the V2 property (e.g., type of null subjects, Tobler-Mussafia) have not yet been systematically exploited. Investigating such phenomena would help us to independently assess the V2 status of these languages and at the same time to clean out the confounding factors related to the interpretation of subject inversion as a consequence of V-to-C or as a vP peripheral phenomenon. The systematic study of word order patterns in a treebank of calibrated texts would allow the respective situation of Medieval Romance languages through time to be assessed (as it was for German and Scandinavian varieties by e.g., Cichosz, Gaszewski and Pezik 2016). This would help us consider the relation between rule and usage, assess whether the observable differential patterns are entirely accounted for by the current models (of a distinction between a strict ForceP and a non-strict FinP V2), and establish when a language loses V2 status.

We present the result of a comparative analysis of V2 word order in two Medieval Romance varieties, French and Venetian (e.g., Wolfe 2018 and references therein). The analysis relies on a calibrated corpus to enhance comparability of results. For each language, the corpus contains one text per century for the 14th, 15th and 16th century, at temporal intervals of about a hundred years, and from the same region for the French material (Normandy). Prose texts are used, belonging to a non-literary legal genre that contain traces of dialogal exchanges, and have been found through preliminary investigations to yield less conservative rates of use of changing variables (Larrivé 2022). The syntactically 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 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).

The extraction process however raises methodological questions, depending on how configurations are counted. Some are well-known, such as the status of a coordinated clauses that under subject coreference is ambiguous between V1 and V2. Some other are less so, such as the surprising high frequency of V1 word order. As both Venetian and French are expected to go from a V2 system to a SVO word order, the high proportions of V1 in the early stages is troubling. A qualitative examination of the V1 order shows that the surprising proportions are found in relative clauses introduced by a subject relative marker and in coordinated clauses, as already expected. Both coordinated clauses and subordinates introduced a subject relative marker are therefore set aside from the count.

What the data reveal is that French is moving from a low FinP V2 to an SVO, while Venetian is one step behind, moving from a high ForceP to a low FinP V2. The data in a) show a consolidation of the verb in the second position in French, due to the rise of the SVO

structure; those in b) show a wide diversification of verb positions in Venetian, due to the increasing availability of Topic projections in front of the inflected V, reflected by the decreasing V2 proportion shown in a); those in c) are a reflex of the loss of null subjects in French and the development of the modern pro-drop system in Venetian.

- a) V2 in Venetian goes down from 70% to 45%, while in French it rises from 40% to 51%
- b) Aggregating V3, V4 and V5, Venetian V3+ rises from 18% to 38%, and French slides from 27% to 18%.
- c) Venetian V1 gradually goes up from 9% to 17%, while in French it decreases from 19% to 9%.

The data in d) show a progressive increase in subject inversion in embedded clauses in both Venetian and French, probably due to the development of vP-peripheral subjects. In Venetian, the raise of subject inversion in main clauses is due to the sum of the new system of vP-peripheral subject while the V-to-C property is still in place, as the steady differential between main and subordinate clauses shows (see e). In French, the low proportion of subject inversion in main clauses shows the loss of V2 (see f)), confirmed by the raising proportion of (vP peripheral) postverbal subjects in embedded clauses (see d)).

- d) The proportion of subject inversion in subordinate clauses raises gradually in both varieties, from less than 10% of all inversions in the earliest period to 40% in the latter.
- e) In Venetian, main-clause subject inversion goes up from 20% in the earlier period to 32% in the two later texts, with a steadily higher rate (+15-18%) with respect to subordinate-clause subject inversion.
- f) In French, main-clause subject inversion stays below 5% across the period, with an increasingly higher proportion of verb in a low position in the later text, as opposed to movement to a higher CP position in the earlier period.

The loss of main/subordinate asymmetry in subject inversion and null subjects is a clear indication that V2 is being lost in both varieties, although the change was already observable in French well before it was in Venetian. A further confirmation of V2 loss in Venetian is the crash of enclisis to inflected verbs (Tobler-Mussafia, see Benincà 2006), highlighted in g), a phenomenon absent in French.

- g) Rate of enclisis to finite Vs in Venetian: \cong 20% until 1340, < 2% after that date.

The significance of the investigation is that the move from V2 to SVO is diagnosed not only by the rates of subject inversion, but also by the rates of pro-drop asymmetry between main and subordinates and the loss of enclisis to the inflected verb in Venetian. This is convergent with an overall narrative by which high verb movement is lost gradually rather than in one go from V2 to SVO, undergoing a phase of movement to a low C position attested by Venetian, where the amount of V3/V* raises but subject inversion and the main/embedded asymmetry remains robust. The systematic investigation of treebank allows to both testing of existing models and expand our understanding of language change, stressing the necessity of a multifactorial analysis by which V2 is acquired and lost through a variety of indicators (and not simply the subject position) and in a stepwise fashion.

BENINCÀ. 2006. A detailed map of the left periphery of Medieval Romance. In Zanuttini, Campos, Herbunger & Portner (eds.), *Crosslinguistic research in syntax and semantics. Negation, tense and clausal architecture*. Washington: Georgetown University Press. 53-86. BELLETTI. 2004. Aspects of the Low IP Area. In Rizzi (ed.) *The Structure of CP and IP: The Cartography of Syntactic Structures Volume 2*. 16-51. CICHOSZ, GASZEWSKI & PEZIK. 2016. *Element Order in Old English and Old High German Translations*. Amsterdam: Benjamins. GROBOL & CRABBE. 2021. Analyse en dépendances du français des plongements contextualisés. In *TALN-RÉCITAL 2021*, https://hal.archives-ouvertes.fr/hal-03223424/file/HOPS_final.pdf. KAISER. 2002. *Verbstellung und Verbstellungswandel in den romanischen Sprachen*. Tübingen: Niemeyer. LARRIVÉE. 2022. Is Medieval French diglossic? New evidence on remnant V2 and register. *Isogloss* 8(2), 1-16. LIU. 2010. Dependency direction as a means of word-order typology: A method based on dependency treebanks. *Lingua* 120(6), 1567-1578. MEISEL, ELSIG & RINKE. 2013. *Language acquisition and change: A morphosyntactic perspective*. Edinburgh: Edinburgh University Press. POLETTI. 2014. *Word order in Old Italian*. Oxford: OUP. RIZZI. 1997. The Fine Structure of the Left Periphery. In Haegemann (ed.), *Elements of Grammar*. Dordrecht-Boston-London: Kluwer academic publishers. 281-337. SITARIDOU, BREITBARTH, BOUZOUITA, DANCKAERT & FARASYN. 2019. Against V2 in Old Spanish. In Breitbarth, Bouzouita, Danckaert &

Farasyn (eds), *The determinants of diachronic stability*. Amsterdam: Benjamins. 131-156 WOLFE. 2018. *Verb Second in Medieval Romance*. Oxford: OUP.