The change in the position of the verb in the history of Portuguese: Subject realization, clitic placement, and prosody

CHARLOTTE GALVES   MARIA CLARA PAIXÃO DE SOUSA

Universidade de Campinas   *Universidade de São Paulo*

This article analyzes the changes in subject position in Portuguese between the sixteenth and nineteenth centuries in terms of the loss of verb-second grammar properties and the rise of an SVO grammar. Our analysis is based on the survey of an unprecedented amount of data for sixteenth-to nineteenth-century Portuguese in a syntactically annotated corpus. We argue that in Classical Portuguese (sixteenth to seventeenth centuries) the verb moves to C(omp), there is no preverbal position reserved for subjects, and all of the preverbal phrases are discourse-prominent constituents—which characterizes Classical Portuguese as a V2-type grammar. In Modern European Portuguese (from the eighteenth century on), in contrast, the verb does not move as high as C(omp), and there is a preverbal position reserved for subjects—in other words, this is an SVO grammar. We suggest that this change from a verb-movement, V2-type grammar to an SVO grammar derived from a prosodic change that happened in the seventeenth century, which also affected clitic placement.*

Keywords: Classical Portuguese, Modern European Portuguese, V2 grammars, V-to-C movement, clitic placement, subject position, prosody-driven language change

1. INTRODUCTION. In this article, we describe the change in subject position and clitic placement in the history of Portuguese between the sixteenth and nineteenth centuries, based on data extracted from sixteen syntactically annotated texts from the Tycho Brahe corpus,[1] written by Portuguese authors born between 1502 and 1836. We propose that the change from Classical to Modern European Portuguese may be interpreted as the loss of the movement of the verb to a position in the left periphery of the sentence that, following tradition, we take to be C(omp). Such a movement is reminiscent of the movement observed in verb-second ('V2') grammars, but with the important difference that the V2 linear order is not always derived. Our analysis is based on the claim that, in Classical Portuguese, the movement of a phrase to the left of the verb is strictly dependent on discourse conditions, and we suggest that the change to Modern Portuguese derived from a prosodic change that happened in the seventeenth century, which also affected clitic placement.

Three points merit general clarification at this initial stage: our stance on the periodization of Portuguese, our stance on V2 and verb movement to Comp ('V-to-C'), and our stance on 'diachronic variation'. Concerning the periodization of Portuguese, what we refer to here as Classical Portuguese (henceforth ClP) is the language instantiated in texts written by authors born between 1500 and 1700, and our focus is on the change from ClP to Modern European Portuguese (henceforth EP, following a long-standing tradition). Previous studies have indicated that ClP corresponds to a grammar that is

[1] The Tycho Brahe Parsed Corpus of Annotated Portuguese consists of texts written by authors born between 1380 and 1836 and currently comprises seventy-six texts (twenty of which are available with syntactic annotation). The corpus is freely available for research and download at http://www.tycho.iel.unicamp.br/~tycho/corpus.

different from both Old Portuguese and EP.[2] ClP is a pro-drop, V-to-C language (see Torres Moraes 1995, Galves 1996, Paixão de Sousa 2004), with pronominal enclisis in V1 clauses and an enclisis/proclisis alternation in nondependent V2 clauses (see Martins 1994, Torres Moraes 1995, Paixão de Sousa 2004, Galves et al. 2005). Old Portuguese was also a pro-drop language with evidence of V-to-C (Ribeiro 1995), but with important differences regarding clitic placement (Namiuti 2008). EP is a pro-drop, SVO language, with pronominal enclisis in V1 and obligatory enclisis in SV affirmative nondependent clauses.

Our analysis of the change from ClP to EP as the loss of V-to-C and the rise of SVO places our proposals within the scope of the much debated issue of the nature of verb-second in Old Romance languages. This, first of all, presents an issue of terminology. In this regard, it is worth pointing out here that the notion of 'V2' in the literature may be found either to refer mainly to linear order, or to be more focused on structural properties. In the former perspective, saying that a language is V2 implies that it excludes or drastically restricts the occurrence of V1 and V>2 (Kaiser 1999, Fieis 2002, Rinke 2009, among others). In the latter perspective, although it does play a role in the discussion, linear order is less crucial in determining whether a language can be classified as V2, and a special emphasis is placed on the movement of the verb to C and on the absence of a preverbal position dedicated to the subject. In this article, we rely on a structural definition of V2, where the crucial property is the movement of V to a high position in the clause. But the presence of a phrase in the preverbal position is optional in ClP; this optionality distinguishes ClP from strict V2 languages (and, in fact, would exclude it from the V2 family in the strict linear-order perspective on V2 mentioned above), in terms to be addressed further on. We consider, however, that except for this optionality, ClP shares important properties with strict V2 languages, mainly with respect to subjects. One of those properties is the absence of a dedicated position for preverbal subjects in V2 languages: like any other initial phrase, preverbal subjects are in topic position; and as a correlated fact, subjects appear in postverbal position more frequently in V2 than in SVO systems. The data for subject positions in our corpus, as is detailed further on, shows this to be the case in ClP. As for postverbal subjects, in V-to-C systems (as opposed to contemporary Romance pro-drop languages) the postverbal position can be a structurally high position, arguably the specifier of IP. As we shall see, the data indicates that this is the case in ClP, as shown by patterns in the position of adverbs and by the difference in the informational status of postverbal subjects in sixteenth- to seventeenth-century texts versus what is observed in EP.

If we are on the right track, the question arises of why the V-to-C properties were lost in the history of Portuguese. This fact is reminiscent of the loss of the same kind of properties observed in the history of other Romance languages—as debated by, among others, Yang (2002). The data on ClP, as we argue below, indicates the need for a review of Yang's (2002) explanation for the loss of V2 as a consequence of its coexistence with null subjects, which would make such grammars 'intrinsically unstable', based on his observations on the diachrony of French. We argue that this is not an adequate account of what happened in Portuguese, since the V2/V-to-C properties had coexisted with null subjects for at least five centuries before the emergence of an SV grammar. Instead, the

---

[2] In the more traditional approach, Old Portuguese is the period that goes from the first remaining manuscripts, at the turn of the twelfth century, to the first half of the sixteenth century. However, many scholars divide it into two periods. The first corresponds to the Galician-Portuguese phase, which lasts until the mid-fourteenth century. The second is a transition period between Galician-Portuguese and ClP, generally called Middle Portuguese (cf. Castro 2006).

loss of V-to-C in ClP is related to the change in clitic placement (Galves et al. 2005), which, in turn, derives from a change in the rhythmic pattern of the language (Galves & Galves 1995). We propose that an increase of enclisis with preverbal subjects related to the new prosodic pattern of the language contributes to the loss of markedness previously associated with this order, and to the nonidentification of preverbal subjects as high topics. This gives greater weight to the reanalysis of preverbal subjects as occupying the canonical preverbal subject position, and no longer a topic position.

As is clear from these initial remarks, the debate in this article is strongly derived from the word-order variation we find in our corpus, and a general comment on our approach to the syntactic variation found in historical texts is now in order. Our approach to the empirical data in our survey follows Kroch (1989, 2001), for whom diachronic variation found in historical texts is the consequence, not the cause, of grammatical change. In other words, the empirical variation is the product of the competition between two coexisting grammars in texts: the older conservative one, on the one hand, and the newer innovative one, on the other, with the gradual replacement of the former by the latter. Diachronic variation, in this view, represents the implementation of syntactic changes. However, it is important to mention that we do find a second kind of variation in historical texts, which does not represent the effect of grammatical change, but rather results from the choice between different constructions that are possible within a grammar—that is, between different options provided by a single grammar in a given period. In this case, as we shall see, the variation occurs mainly between authors, who may choose the different options at varying rates. But it can also discriminate groups of authors over time. The difference in nature of the two kinds of variation can be argued on both quantitative and qualitative grounds. Diachronic variation, on the one hand, presents an S-shaped curve (Kroch 1989) and is associated with qualitative changes. Synchronic variation, on the other, presents neither a clear quantitative diachronic trend nor all of the qualitative signs of grammatical change, as can be shown by the stable presence of other syntactic phenomena, which strongly suggest that the grammar is the same despite the variation found in some constructions. Our data on sixteenth- to-nineteenth-century Portuguese presents the two kinds of variation, and we interpret their relevance to our analysis in different ways, in accordance with the approach briefly outlined above.

The article is organized as follows. We first present newly surveyed data and explore aspects of subject-verb order in ClP (§2), which we take as essential for understanding the grammatical change. We then analyze the dynamics of the change and propose an interpretation for the loss of V-to-C in ClP and its relation with pro-drop in §3. We argue that, contrary to what happens in Germanic languages, V-to-C in pro-drop languages is entirely dependent on discourse conditions. Therefore, it is more prone to suffer the effects of prosodic changes. Finally, in the concluding remarks, we point out some implications of our proposal for the analysis of the difference between Germanic and Romance languages, and for the notion of competing grammars.

2. Subject position in the diachrony of portuguese. The survey to be presented in this article was conducted with an unprecedented amount of data for sixteenth- to nineteenth-century Portuguese, comprising 34,293 tokens in texts written by sixteen authors born between 1502 and 1836.[3] This data has been extracted from a syn-

---

[3] The parsed corpus is composed of the following authors and works: Pero Magalhães Gândavo (b. 1502), *Historia da Província de Santa Cruz vulgarmente chamada Brasil*; Fernão Mendes Pinto (b. 1510), *Peregrinação*; Diogo do Couto (b. 1542), *Décadas*; Frei Luís de Sousa (b. 1556), *A vida de Frei Bertolameu dos*

tactically annotated corpus and represents the first results made possible by this annotation. Some of the aspects to be discussed in this section have been partially tackled in previous studies—in particular, Paixão de Sousa 2004, Galves et al. 2005, and Galves & Paixão de Sousa 2005—all of which, however, were conducted without the benefit of syntactic annotation and were limited to word order in sentences with clitic pronouns. The present article confirms the results of these previous works and discusses sentences both with and without clitics, thus broadening the universe of analysis.

As we argue further below, we interpret the results of this survey as important evidence of a V2/V-to-C system in the authors of the first phase (i.e. sixteenth- and seventeenth-century authors). As with other Old Romance languages, however, 'V2' in ClP must be understood in a particular context: that of a null-subject system. Among other effects, this means that in this language superficial verb-second coexists liberally with verb-first. Our central aim in this article is to model the change between such a system and a canonical SV system, which is what is found in EP. In accordance with this aim, we focus this section on the possible expressions of subjects in ClP: lexical (postverbal or preverbal) and null. Moreover, the various possible positions for the verb will be shown with reference to the data and discussion presented by Cavalcante and colleagues (2010), Gibrail (2010), and Antonelli (2011).

2.1. PREVERBAL, POSTVERBAL, AND NULL SUBJECT FREQUENCY RATES. The most remarkable fact about subject positions in ClP is the high frequency of constructions with postverbal lexical subjects versus constructions with preverbal lexical subjects, as compared to EP. Figure 1 shows the rates of '(X)VS' (constructions with postverbal subjects, including verb-initial constructions and constructions with a constituent other than the subject before the verb) over 'SV(X)' in sixteenth- to nineteenth-century texts.[4]

The data depicted in Fig. 1 reveals two clear groups: in texts written by authors born in the sixteenth and seventeenth centuries, postverbal subjects range from 36% to 77%,[5]

Mártires; Manuel Galhegos (b. 1597), *Gazeta*; Pe Antonio Vieira (b. 1608), *Sermões*; Maria do Céu (b. 1658), *Vida e morte de Madre Elena da Cruz*; André de Barros (b. 1675), *Vida do apostólico Pe Antonio Vieira*; Cavaleiro de Oliveira (b. 1702), *Cartas*; Matias Aires (b. 1705), *Reflexões sobre a vaidade dos homens*; Marquesa de Alorna (b. 1750), *Cartas*; J. D. Rodrigues da Costa (b. 1757), *Entremezes de cordel*; Almeida Garrett (b. 1799), *Teatro*; Marquês de Fronteira e Alorna (b. 1802), *Memórias*; Camilo Castelo Branco (b. 1825), *Maria Moisés*; Ramalho Ortigão (b. 1836), *Cartas a Emília*. For full references on each text, please refer to http://www.tycho.iel.unicamp.br/~tycho/corpus.

The data were extracted from 19,850 nondependent clauses and included all verbs except *ser* 'to be'. Because they are categorically VS, we dropped from the quantification all occurrences of parenthetical clauses following direct speech, as in (i).

(i) Como?   – **exclamou António de Queirós.**
    How.come?   – **shouted   Antonio de Queirós**.

[4] In what follows, the date assigned to each text is the birthdate of its author, for both a theoretical reason and a pragmatic reason. As for the former, it is consistent with the generativist claim that speakers select their grammars in the process of first language acquisition and these do not change in individuals after the critical period. The pragmatic reason has to do with the fact that in some cases it is difficult to date the texts themselves, either because they were published long after they were written or because they are composed of pieces written over a long period of time. This is the case, for instance, of correspondence texts. Paixão de Sousa (2004:200–207) discusses this issue at length and argues that the dynamics of change make more sense if we refer to the birth of the authors than if we use the date of the texts.

[5] There is in our corpus a remarkably elevated use of VS in the texts written by the last generation of authors born in the seventeenth century. According to Paixão de Sousa (2004), this characteristic of the texts produced by this generation is connected to a stylistic tendency, as the last 'baroque' authors sought to enhance the features present in the writings of the former, more canonical generation. As is seen in more detail in Table 1 and Fig. 2 below, it is interesting to note that the last author of the 1500–1700 generation, André de
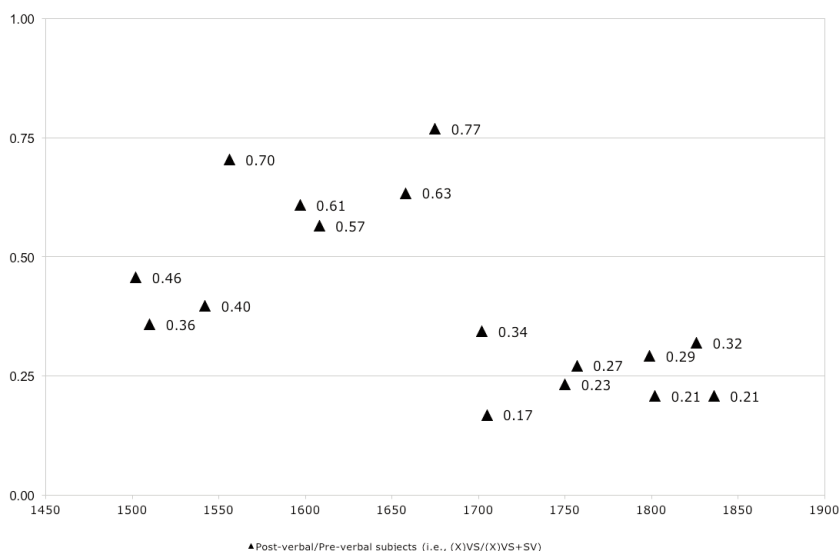
FIGURE 1. The evolution of the rate of (X)VS in matrix clauses with lexical subjects,
sixteenth to nineteenth centuries.

whereas in texts written by authors born after 1700, postverbal subjects range from 17%
to 34%. Remarkably, while the relative rates of postverbal/preverbal subjects are higher
in the seventeenth century than in the sixteenth century (right before a steep decrease in
the eighteenth century), when the whole sixteenth- to nineteenth-century period is taken
into account, the gradual tendency for the proportion of constructions with postverbal
subjects to fall is statistically significant.[6]

　　We take the high relative frequency of postverbal subjects in sixteenth- to seven-
teenth-century texts as one of the most important pieces of evidence for a V-to-C system
in ClP. Other important evidence for our analysis, apart from the position of subjects it-
self, is the presence of nonsubject preverbal constituents; in what follows, we analyze
not only constructions with postverbal subjects and nonsubject fronted constituents
('XVS', included in the data in Fig. 1), but also constructions with null subjects and
nonsubject fronted constituents (not yet considered in Fig. 1, but to be tackled below)—
in other words, all constructions in which a constituent other than the subject occupies
the preverbal position, that is, 'non-SV' orders.[7] We claim that in ClP, preverbal posi-

Barros, presents the text with the highest general rate of VS of all, 41%—and, again according to Paixão de
Sousa (2004), the late 'baroque' style is particularly noticeable in Barros, according to independent literary
studies. For the purposes of this article, there are two relevant issues at stake in this regard. The first point is
that, whatever stylistic options were chosen by the late seventeenth-century authors, those options were only
possible within the restrictions of their grammar—which, as we see it, allowed for flexible order conditioned
by discourse (and thus a wider variation regarding SV/VS). The second point is that irrespective of the liter-
ary styles in vogue for the generations born after 1700, the rate of postverbal subjects never rose again above
8–12% in their texts.

　　[6] Logistic regression on the data set shows that the usage rate of VS (in contrast with SV) drops signifi-
cantly between the sixteenth- and nineteenth-century authors (with a log-odds of −0.005 per year, $p < 0.001$,
or more precisely, $< 2e\text{-}16$). Regression was applied with Language Variation Suite (LVS), available at
https://languagevariationsuite.shinyapps.io/Pages/.

　　[7] A referee raises the issue of a possible shift in the interpretation associated with V1 before and after the
change. We leave this very interesting question for further research. What we can say is that the frequency of
V1 decreases after 1700 (see Figs. 5 and 6 below), although there is no difference in the frequency of null sub-
jects between the seventeenth and the eighteenth centuries (Fig. 2).

tion is reserved for discourse-prominent elements. We are not concerned, in this article, with the exact nature of this position (and we refer to Antonelli 2011 for a concrete proposal in the framework of Rizzi's cartography (e.g. Rizzi 1997)). The relevant point is that ClP is not an SV system, and preverbal subjects occupy the same position as other fronted elements.

In 1–3 below, we illustrate all of the possible patterns for the expression of the subject in the corpus: preverbal (1), postverbal (2), and null (3).[8]

(1) **Christo Senhor nosso**, **disse**   a   seus Discipulos, que o   segredo
     Christ Lord   our   say.3SG.PST to his disciples   that the secret
          d'aquelle dia é reservado só   ao   Padre.
          of:that   day is reserved only to.the Father
     'Christ our Lord said to his Disciples that the secret of that day is reserved
       only to the Father.'   (Vieira, 1608)

(2) **Começou   el-Rei** a   igreja de São Vicente.
     start.3SG.PST the-king the church of São Vicente
          'The king started the church of São Vicente.'   (Sousa, 1556)

(3) e   com a   sua prisão   **mudaram**   de intento
     and with the his imprisonment change.3PL.PST of intent
          'and with his imprisonment, they changed their intent'   (Galhegos, 1597)

The frequency of each of these three possibilities for the expression of subjects in ClP is interestingly different from what might be documented for a canonical pro-drop SV system (such as EP). Figure 2 shows the distribution of the possible patterns for subjects in main clauses, illustrated in 1, 2, and 3 above, between 1500 and 1850.
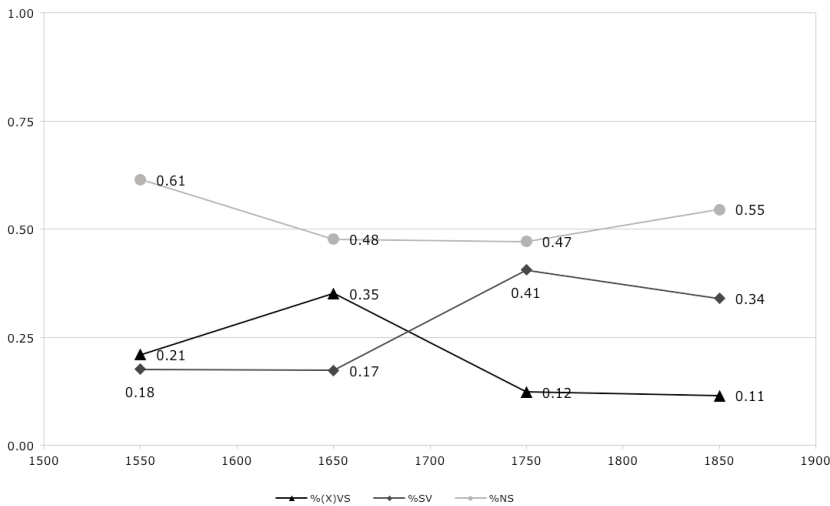

FIGURE 2. The expression of subjects in main clauses, sixteenth to nineteenth centuries, by 100-year periods (proportions of NS: null subjects, SV: preverbal subjects, (X)VS: postverbal subjects).

As we can see, in sixteenth- and seventeenth-century texts, respectively, the proportion of postverbal subjects is 21% and 35%, while the proportion of preverbal subjects is 18% and 17%, and null subjects represent 61% and 48% of main clauses. In eigh-

[8] The following abbreviations are used: 1/2/3: 1st/2nd/3rd person, CL: clitic, DAT: dative, FUT: future, INF: infinitive, PL: plural, PRS: present, PST: past, REFL: reflexive, SG: singular.

teenth- and nineteenth-century texts, respectively, the proportion of postverbal subjects is 12% and 11%, and the proportion of preverbal subjects is 41% and 34%. In short, before 1700 postverbal subjects are more frequent than preverbal subjects (and less frequent than null subjects); after 1700, postverbal subjects are less frequent than preverbal subjects (and still less frequent than null subjects). In other words, it is only after 1700 that preverbal becomes a favored position for lexical subjects in Portuguese texts. Consider, in particular, that because we are comparing all of the possibilities for subject expression (including null), we can clearly see that the decrease in postverbal subjects at the turn of the eighteenth century in fact corresponds to an increase in preverbal subjects (and not, for example, to an increase in the rate of null subjects, although this rate may oscillate; see below).[9] To sum up:

(i) The frequency of postverbal subjects decreases between the seventeenth and the eighteenth centuries.
(ii) The frequency of preverbal subjects surpasses that of postverbal subjects at this same point.
(iii) Null-subject proportions vary, but show no clear tendency.

Although this general picture very much encompasses the fundamental data to be discussed in §3, it is fair to point out that there is interesting variation among contemporary texts in the corpus and that this variation also contrasts the two periods under study (pre-eighteenth century and post-eighteenth century). Table 1 and Figure 3 show this, detailing the data for postverbal, preverbal, and null subjects previously shown above, this time for each of the sixteen texts of our corpus.

		NS		SV		(X)VS	
		N	*%*	*N*	*%*	*N*	*%*
1502	Gandavo	378/575	66	107/575	19	90/575	16
1510	Pinto	554/816	68	168/816	21	94/816	12
1542	Couto	572/947	60	226/947	24	149/947	16
1556	Sousa	1,049/1,539	68	145/1,539	9	345/1,539	22
1597	Galhegos	539/1,153	47	240/1,153	21	374/1,153	32
1608	Vieira	355/942	38	255/942	27	332/942	35
1658	Céu	541/904	60	133/904	15	230/904	25
1675	Barros	653/1,405	46	174/1,405	12	578/1,405	41
1702	Caval.	956/1,676	57	473/1,676	28	247/1,676	15
1705	Aires	813/2,298	35	1,235/2,298	54	250/2,298	11
1750	Marq.	449/850	53	308/850	36	93/850	11
1757	Costa	486/913	53	311/913	34	116/913	13
1799	Garrett	1,038/1,709	61	475/1,709	28	196/1,709	11
1802	Alorna	833/1,754	47	730/1,754	42	191/1,754	11
1826	Camilo	561/1,119	50	380/1,119	34	178/1,119	16
1836	Ortigão	749/1,250	60	397/1,250	32	104/1,250	8

TABLE 1. Breakdown of the expression of subjects in main clauses by author, sixteenth to nineteenth centuries (NS: null subjects, SV: preverbal subjects, (X)VS: postverbal subjects).

The rates of null subjects vary considerably throughout the corpus (ranging, irregularly, from 35% in the text by Aires to 68% in the texts by Pinto and Sousa). We con-

---

[9] Logistic regression on this data set (i.e. considering all possibilities: SV, VS, and null subject) shows that the usage rate of VS (in proportion with SV and null subjects) drops significantly between the sixteenth- and nineteenth-century authors (with a log-odds of −0.001 per year, *p* < 0.001, or more precisely, < 2.2e-16). In contrast, the usage rate of SV (in proportion with VS and null subjects) rises significantly between the sixteenth- and nineteenth-century authors (with a log-odds of 0.002 per year, *p* < 0.001, or more precisely, < 2.2e-16). Regression was again applied with LVS (see n. 6).
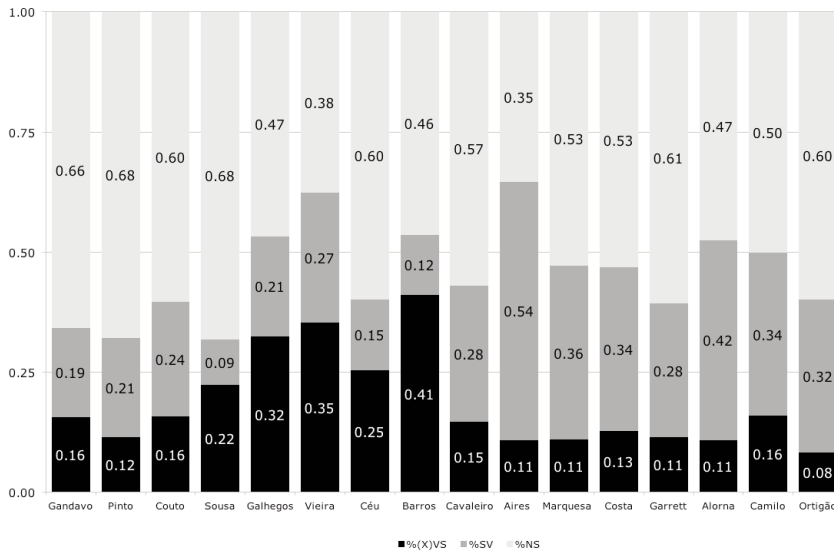
FIGURE 3. Graphical representation of the expression of subjects in main clauses by author, sixteenth to nineteenth centuries (% null subjects, % SV, % (X)VS).

sider this ample variation to be expected, as the option for null subjects in a pro-drop language is related to textual and not strictly grammatical constraints. But most importantly, we do not consider the variation in the incidence of null subjects in Portuguese texts throughout this period to indicate parametric change as far as the null-subject parameter is concerned.

We are more interested here in the variations concerning the position of lexical subjects. This, of course, will also be related to textual constraints—but the effects of such constraints will be different in different grammars (a point to which we return more precisely below). Crucially, in our data the variation in the position of lexical subjects presents different patterns in different subperiods, and we interpret this as indicating a frontier between periods representing different grammars. As Fig. 3 shows, sixteenth- to seventeenth-century authors present a wide range of variation in the position of lexical subjects: (X)VS ranges from 12% to 41%, and SV from 9% to 27%. Additionally, within this period we observe a higher rate of (X)VS versus SV(X) in the authors born after 1550. However, for eighteenth- to nineteenth-century authors, a more consistent pattern emerges, in which the (X)VS range of variation is narrower: from 8% to 16%. Furthermore, for all eighteenth- to nineteenth-century authors, SV(X) rates are consistently higher than the rates of (X)VS (see the rates of SV(X) versus (X)VS in Table 3 and Fig. 3 for each text after Cavaleiro, born in 1702). This is not true for sixteenth- to seventeenth-century authors: in this period, some authors present more SV(X) than (X)VS, and others more (X)VS than SV(X) (as can be seen in Fig. 3 for every text before Barros, born in 1675).

Our first conclusion from the data presented so far is that there are two distinct patterns of variation in the position of lexical subjects over the period included in the corpus: the variation in sixteenth- to seventeenth-century texts occurs within a wider margin and shows more contrast across different authors, whereas the variation in eighteenth- to nineteenth-century texts occurs within a narrower margin and is more consistent across authors. As for the sixteenth- to seventeenth-century period, we show in §3.2 that, although the first sixteenth-century texts in the corpus present a higher rate of

SV(X) than of (X)VS, these texts pattern with the remaining sixteenth- to seventeenth-century texts with respect to their frequency of XV in general (i.e. including XV with null subjects). Crucially, it is only after the eighteenth century that we see SV(X) as a clearly predominant pattern in the texts (a central fact to which we return below).

Further evidence on the position of subjects comes from Galves and Gibrail (2018), who study the syntax of subjects in both matrix and subordinate clauses in the same corpus, focusing on transitive verbs. In this study, the authors show that postverbal subjects are much less frequent in subordinate clauses than in matrix clauses in sixteenth- to seventeenth-century texts, which is to be expected if VS derives from the movement of V to C. This can be seen in Figure 4, which shows that while the frequency of VSO in matrix clauses is 33% in the sixteenth century and 46% in the seventeenth century, in subordinate clauses it is only 15% and 19%, respectively (note, incidentally, that this difference is no longer observed after 1700). It is worth noting, moreover, that these VSO cases in subordinate clauses are cases in which either the complementizer is null or the main verb allows for CP recursion, which in both cases gives room for V-to-C, and consequently VS order.[10]
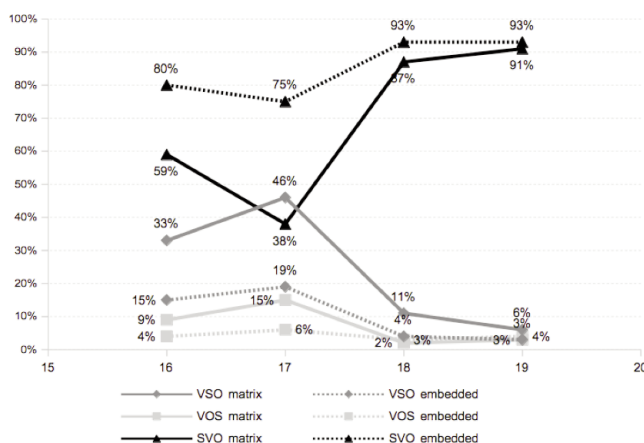

FIGURE 4. SVO, VSO, and VOS in matrix and embedded clauses, by century
(to appear in Galves & Gibrail 2018).

In this same study, Galves and Gibrail (2018) also find several instances of the VSAdv order, as shown in example 4.

(4) e　　sendo este egócio dos　　maiores, que podia ter　　uma Monarquia,
　　and being this deal　　of.the greatest that could have a　　monarchy
　　　　o　　　　　　 **fiou**　　　　　 **o animoso Rei inteiramente** da
　　　　3SG.ACC.CL entrust.3SG.PST the brave　　king entirely　　of.the
　　　　intelligência e　　indústri de um só　　homem
　　　　intelligence and industry of a　　only man
　　　　'and being among the greatest of deals which a Monarchy could have, the
　　　　brave king entrusted it entirely to the intelligence and industry of just
　　　　one man'　　　　　　　　　　　　　　　　　　　　　　　　　(Barros, 1675)

Orders such as in 4, with the adverb following the postverbal subject, are typical of Germanic languages, in contrast with what is expected in Romance, where adverbs precede

---

[10] As for VOS, which is rather marginal, both in matrix and subordinate clauses, Galves and Gibrail (2018) show evidence that it is derived by VO movement, which they claim targets the C layer.

the preverbal subject, as Belletti (2004) has shown. In combination with the data shown above concerning embedded clauses, we interpret this as further evidence that VS in ClP is correlated with V-to-C movement.

In §2.2, we turn to another important perspective on the data debated in this article: the interpretation of subjects according to their position.

2.2. THE INTERPRETATION OF PREVERBAL AND POSTVERBAL SUBJECTS. A second important point we wish to make here with regard to the data on subjects in sixteenth- to seventeenth-century texts is that the preverbal position in ClP is reserved for prominent elements of varied kinds, all of them dependent on information structure.[11] In the case of written texts (which is what we have in our hands), we may consider the text itself as the discourse context—which means that different texts will produce different patterns of word orders. Paixão de Sousa (2004) pointed out, for instance, that SV(X) is particularly frequent in texts with abundant narrative sequences, in which different participants in the narrated event take their turns in a segment of the text. This is the case of the text *Peregrinação*, by Fernão Mendes Pinto (b. 1510), included in the present study, as illustrated in the following examples.

(5) **O capitão-mor lhe   respondeu**   que os embaixadores
    the captain-major 3SG.DAT.CL answer.3SG.PST that the ambassadors
        tinham   seguro   para suas pessoas.
        have.3PL.PST insurance for   their persons
      'The captain-major answered him that the ambassadors had personal insurance.'   (Pinto, 1510)

(6) **O brâmene lhe   deu**   por isso seus agradecimentos
    the brahman 3SG.DAT.CL give.3SG.PST for that his   thanks
      'The brahman thanked him for that'   (Pinto, 1510)

(7) **nosso capitão-mor cometeu**   então queimar=lhe   a galé
    our   captain-major commit.3SG.PST then   burn.INF=3SG.DAT.CL the galley
      'our captain-major proceeded then to burn his galley'   (Pinto, 1510)

As becomes clear when one examines the context from which the sentences above have been extracted, they are part of a narrative sequence that may be summarized as an account of a discussion between two characters: 'the captain' and 'the brahman'. Each time one of the characters is mentioned, describing their turn in the conversation, a constituent that refers to that character occupies the preverbal position, that is, the normal position for V2-topics.

The same phrasal organization can be seen, naturally, in sequences in which the fronted element is not the subject. In such cases, we have XV(S) sequences—that is, XV, with null subjects; or XVS, with postverbal lexical subjects. Take, for instance, the following sentences from the same text.

(8) Ao   mercador que me   trouxe   **mandou   Pero de Faria**
    to.the merchant that 1SG.ACC.CL bring.3SG.PST order.3SG.PST Pero de Faria
        dar   sessenta cruzados ...
        give.INF sixty   cruzados ...
      'To the merchant who brought me, Pero de Faria ordered that sixty cruzados be given … '   (Pinto, 1510)

---

[11] The exact nature of the feature(s) that trigger(s) movement has been the object of much debate, and is not addressed at this point. We come back to this issue in n. 22 below.

In this case we are again faced with a narrative sequence, with the characters taking turns in the narrated events; but now, the alternating characters (one of whom is 'the merchant' in 8) take turns in receiving presents and instructions from another, constant character (Pero de Faria). Naturally, the grammatical subject of the sentences here corresponds to the 'agent' (*Pero de Faria*), and the dative PP corresponds to the 'target' (*ao mercador que me trouxe*). Because the PP is the prominent constituent in the discourse, it appears preverbally, while the subject, being less prominent, does not.

XVS can therefore be accounted for in ClP as instances of V2-fronting of nonsubjects. The proportion of SV(X) and (X)VS may therefore oscillate considerably according to the discourse nature of the texts in which they appear. It will be crucial for us to show that this oscillation, which is dependent on discourse requirements, is expected to be greater in a V2/V-to-C system than in an SV/non-V-to-C system. As shown above in Table 1 and Fig. 3, this is precisely what happens: post-eighteenth-century texts present less variation among contemporaries than pre-eighteenth-century texts.

Further support for the claim that the interpretation of subjects differs in ClP and EP is found in Galves & Gibrail 2018, the same study referred to above, regarding postverbal subjects in particular. The authors argue that the interpretation of postverbal subjects in the texts from the classical period is different from the interpretation associated with VS in EP—where, according to Costa (2004:79–80), VSO is felicitous only if both the subject and the object are foci, that is, answers to the question 'Who did what?', and SVO is the obligatory order when the whole sentence is new information or only the object is focused. This is clearly not the case in ClP, in which VSO does not require that the subject be the new information of the sentence, and SVO is not required when the entire sentence is new information ('out of the blue'). The two following sentences, from Galhegos (b. 1597), illustrate both cases. In 9, which occurs in the initial position of a narrative, the whole sentence is interpreted as new information, and in 10, where the referent of the subject was previously introduced as one of the protagonists of the story, only the object is focused. In both cases, SVO would be required in EP.

   (9) Em várias   partes das   fronteiras **fizeram**   **os castelhanos** fumo.
         in   several parts of.the borders   make.3PL.PST the Castilians   smoke
           'The Castilians made fire on several parts of the border.'
  (10) **tomaram=lhe**   **os nossos** algumas armas, e   munições
         take.3PL.PST=3SG.DAT.CL the ours   some   weapons and bullets
           'our people took from them some weapons and bullets'

Galves and Gibrail (2018) argue that in most cases, postverbal subjects are interpreted either as familiar topics or as continuing topics in ClP. According to them, 'familiar topics are either proper names which refer to the main characters of the narratives, God and other religious entities (*the Devil*, *the Holy Spirit*), frequently quoted authors (the authors of the Gospels in the Sermons), or abstract or generic entities like "the enemies", "humankind", vices and virtues, etc.'. They give as an example the fact that in the biography of Frei Bertalomeu dos Mártires, written by Frei Luis de Sousa (born in 1556), out of a total of forty postverbal subjects, twenty refer to the Frei, who is the main character of the work. They argue that all of these facts support the hypothesis that postverbal subjects are not in a position associated with focus interpretation—that is, differently from what is argued by Costa (2004) for EP, they are not in a low position in the clause. This is straightforwardly derived if the verb is itself in a high position.

2.3. SUMMING UP. In broad terms, the analysis proposed so far is that SV(X) in classical texts (1500–1600) and SV(X) in modern texts (1700–1800) correspond to different structures and different grammars: in the classical texts, SV(X) corresponds to

constructions in which the preverbal subject is topicalized just like any XV(S) construction. In the modern texts, however, SV(X) corresponds to subjects in canonical positions. In other words, an SVO grammar has emerged. Our central question now is why the old grammar was replaced by the new; this is what we explore in §3 below, following a brief debate on some controversial issues involved in the loss of V2 in Romance in general.

3. Romance V2 revisited. In this section, we propose an account of the loss of the verb-second properties in Portuguese. Contrary to what has been proposed by important work on similar processes, we argue that this change is not related to any form of 'instability' derived from the presence of either null subjects or V1 and V3 orders in ClP. Based chiefly on the patterns of clitic placement, and following Galves and Paixão de Sousa (2005), we argue that the V3 order in this language derives from the availability of (at least) two topic positions preceding the verb, associated with distinct prosodic patterns. Speakers of ClP could differentiate perfectly either position when confronted with superficial V3 or superficial V2, and there was no instability in the system. The piece missing from this puzzle, as we argue later, is prosody.

3.1. Romance V2 and its challenges. In order to discuss this, we first present a summarized account of Yang's (2002) model of language change and his discussion on Romance V2. The model of language change proposed in Yang 2002 derives from a model of acquisition viewed as grammar competition such that, when a grammar successfully analyzes an utterance the child is exposed to, it is rewarded; when it fails, it is punished. More formally speaking: 'Upon the presentation of an input datum s, the child (a) selects a grammar G_i with the probability p_i; (b) analyzes s with G_i; (c) if successful, reward[s] G_i by increasing p_i; otherwise, punish[es] G_i by decreasing p_i' (Yang 2002:26–27). At each step of the acquisition process, therefore, 'each grammar G_i is paired with a weight p_i, which can be viewed as the measure of prominence of G_i in the learner's language faculty' (p. 26). In cases where the linguistic environment is homogeneous, all linguistic expressions are generated by a grammar G_i, and at the end of the process, G_i is the only grammar the learner has access to. In the case of heterogeneous environments, however, no grammar is 100% compatible with the data. In this case, the model allows language learners to 'form internal representations of coexisting grammars' (Yang 2000:241). This feature of the model strongly distinguishes it from other similar proposals of grammar selection, which assume that in case of ambiguity just one grammar is selected, as a result of general principles of economy that lead children to choose the simplest one (see for instance Clark & Roberts 1993, Roberts 2007).

The model has important consequences for grammar change, since, as acknowledged by Yang (2002:33), while a combination of two grammars is synchronically stable, it may be diachronically unstable. This is because, over time, the relative weight of the two grammars may change, and one of them may win the competition as it drives the weight of the other to zero. In fact, the model predicts that 'once a grammar is on the rise, it is unstoppable' (Yang 2002:132).

As a case study, Yang proposes an analysis of the loss of V2 in French. His model requires that in order for an SVO grammar to overtake a V2 grammar, there must be more evidence incompatible with V2 than with SVO. According to him, evidence for V2 is VS (XVSO, OVS), and evidence for SVO is V3 (SXVO, XSVO). The SVO order is ambiguous, however, since it can be generated by both a V2 and an SVO grammar. Yang acknowledges that, from a quantitative point of view, SVO is very frequent in V2 languages, while VS patterns occur at a rate of approximately 25–30%. This means that

the nonambiguous evidence that acquirers of a V2 language have at their disposal has a weight corresponding to this rate. By contrast, in SVO languages (like English) only 10% of all sentences are V3 patterns—that is, are nonambiguous. Following Yang's point of view, this means that the weight of the nonambiguous evidence for SVO is weaker than the weight of the nonambiguous evidence for V2. The competition model therefore predicts that 'the 10% advantage of a SVO grammar cannot throw off a V2 grammar, which has 30% of VS pattern to counter' (Yang 2002:135).[12] This would account for the strong stability of V2 in Germanic languages, but leaves unexplained why V2 could be lost in Old French, and more generally in Old Romance languages. Yang's answer to this puzzle lies in the fact that Old French—as well as the other Old Romance languages—is a null-subject language. In such languages, according to him, the evidence for V2 will be reduced in the total data set because of the existence of null-subject sentences. He then observes that in three texts written in French at the turn of the fifteenth century, while null subjects occur at a rate of c. 40% and V>2 ranges from 11% to 15%, VS ranges from 5% to 18%—quite far from the 30% found in non-pro-drop V2 languages. In the competition model, this accounts for the fact that the SVO + null-subject grammar eventually wins over the Old French V2 grammar. Yang concludes that the combination of pro-drop and V2 is 'intrinsically unstable and will necessarily give away [sic] to an SVO (plus pro-drop) grammar' (2002:137).

This is the statement we wish to debate, by examining whether Yang's model of competition of grammars accounts for the dynamics of the loss of V2 (V-to-C) in the history of another pro-drop Romance language, European Portuguese.

Our data on ClP have so far shown evidence of an increase in preverbal subjects between the seventeenth and eighteenth centuries, as we showed in §2 (see Fig. 2). This would mark the 1700s as the point at which SVO supersedes V2/V-to-C in the diachrony of Portuguese. However, this change in the data is not preceded by a situation where VS is superseded by V3. As shown in Figure 5 (adapted from Cavalcante et al. 2010), ClP texts from the sixteenth and seventeenth centuries present rates of 13% and 10% of V3 sentences, whereas VS patterns in those centuries range from 21% to 35% (cf. Fig. 2).
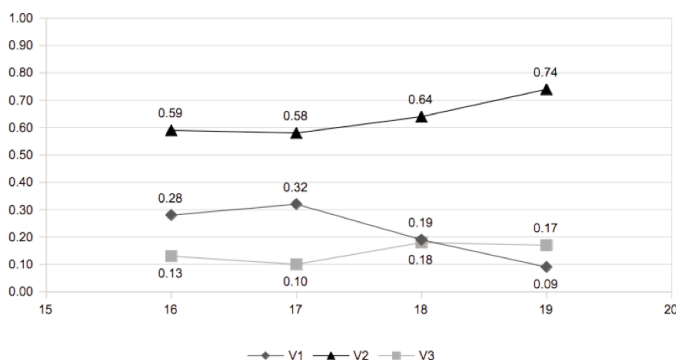

FIGURE 5. Evolution of V1, V2, and V3 in the diachrony of Portuguese, by century
(adapted from Cavalcante et al. 2010).

---

[12] Note also that the same reasoning would lead one to expect that in a heterogeneous environment composed of V2 and SVO speakers, the V2 language would take over the SVO language, since acquirers would end up with more nonambiguous evidence for V2 than for SVO. This may have been the case when Germanic tribes invaded the Roman Empire in a period in which Proto-Romance languages were already SVO.

Notice that from the point of view of Yang's analysis, the advantage of V2 grammars over SVO grammars would be computed from the difference between the frequency of VS and the frequency of V3. Therefore, if we reason within Yang's framework, we meet a puzzling fact: it would seem that in the diachrony of Portuguese, SVO patterns replaced V2 patterns (as shown by Fig. 2 above, as well as by Fig. 6 in §3.2 below), even without a change in the relation between V3 and VS—since the rise in SV is not preceded by a significant change in the rates of V3. Still following Yang's model, this would mean that the V2 grammar was strongly evidenced in the data, and there was no reason why children should have selected the SVO grammar.[13]

The fact remains that the word-order patterns show a steep change at the turn of the 1700s, and this drastic change in the space of one generation must be explained. We attempt to do so below.

3.2. AN ALTERNATIVE ACCOUNT OF ROMANCE V2. Our hypothesis about the loss of V2 in Portuguese depends on a deeper reflection on the difference between a language like ClP and Germanic languages. This reflection profits greatly from the analysis proposed by Frey (2006) for V2 in German and recently discussed and extended to other Germanic languages by Light (2012).

Frey (2006) argues that in German, two different processes underlie the movement of a phrase to Spec/CP, yielding V2 order. One is a formal movement, due to the formal (extended projection principle; EPP) properties of Comp, more specifically Fin, and the other is the topicalization of some element of the clause, which Frey calls 'True A-bar-movement (TAB)', to higher positions in the CP field. Frey argues at length that only A-bar movement entails a pragmatically marked interpretation, which is, according to him, one of contrast.[14] This accounts for the long-standing fact that the interpretation associated with the preverbal position in German is not the same in all cases, which would remain unexplained under a unitary analysis of movement to Spec/CP. This also accounts for the fact that SVO is by far the most common order in German (c. 70% according to several authors; cf. Lightfoot 1997 quoted in Frey 2006, among others). Since subjects occupy the highest position in the so-called MIDDLE FIELD (the post-Fin layer of the sentence), they are the more natural candidates to undergo movement in order to satisfy the EPP feature of Fin. According to Frey, however, it is not just the subject that can be moved by FORMAL MOVEMENT. Any other constituent moved higher than the subject by SCRAMBLING in the middle field can undergo formal movement as well, and therefore lack contrastive interpretation. As emphasized by Frey and by Light (2012), formal movement is a property typical of V2 languages, and it is satisfied by any constituent occupying the highest position below Fin. Light extends Frey's analysis to other Germanic languages and claims that 'TAB remains a constant across the Germanic language family, but only V2 languages have [formal movement] as an option, because [formal movement] is a phenomenon inherently linked to the V2 requirement, existing solely to fill Spec,CP as a last resort due to what has been described as an EPP feature on C' (Light 2012:vi).

---

[13] In Yang's framework, one could possibly argue that Romance V2 languages are a stable combination of V2 with SVO grammars. Such a claim, however, would leave unexplained why, at some point, instability arises. Instead, we argue in §3.2 that V3 is compatible with the V2 grammar instantiated by ClP, and possibly other Romance languages.

[14] Frey additionally proposes that some adverbs can be inserted directly in Spec/CP. This is not relevant for our analysis.

Attributing both formal movement and TAB to a feature F, Light expresses the difference between English and Germanic V2 languages in the following way:

> In non-V2 languages, such as English, Spec,CP in matrix clauses lacks the EPP requirement which leads to Formal Movement. Therefore, the [F] feature on C does not behave in the same way. Just as in V2 languages, C may carry the requirement to probe for a [+F] feature, and in this case TAB may result, just as before. The difference, however, is that if the [F] feature on C does not find an [F]-marked constituent (a constituent with the [+F] feature), the derivation would crash. When an XP with the [+F] feature is not fronted, no other type of fronting will occur. Thus, in non-V2 languages, we must also assume that a version of C exists which does not carry an [F] feature which must be valued. This is the version of C in the default clause structure of English, where no fronting occurs at all.
>
> This is a rough sketch of how we may account for the relationship between TAB in V2 and non-V2 Germanic languages. Although True A-Bar Movement has clearly been 'adopted,' for lack of a better term, to help satisfy the EPP-like requirement on Spec,CP in V2 languages, any analysis must capture the fact that it exists independent of this requirement, as languages like English clearly show. (Light 2012:155–56)

The idea that we want to put forth here is that TAB is not restricted to Germanic languages, and many other languages have it. As for Romance languages like ClP, we suggest that the similarity with, and the difference from, Germanic V2 languages may lie in the following point: ClP HAS V-to-C (Fin)[15] but LACKS FORMAL MOVEMENT to Spec/C (Fin). Constituent fronting therefore only occurs through TAB, under specific discursive conditions.[16] This would straightforwardly account for the existence of V1. As for the frequency of V3, it suggests that while only one specifier position is licensed at the CP level in V2 languages, more than one is legitimate in languages like ClP. This could well be due to the different position of V. We comment on this briefly in the concluding remarks. Moreover, we shall see in §3.3 below that evidence for two available preverbal positions in ClP comes from clitic placement.

In addition, and very interestingly, this analysis also explains a fact that up to now had remained unnoticed about V2 orders themselves: the rate of XV orders in which X is not a subject is higher in ClP than what is reported for German.[17] It is important to observe that when we compare V2 orders in German and ClP, we must keep in mind that one of them is a null-subject language. So, even if the piece of data presented above (cf. Fig. 2) does show that the frequency of VS is globally higher than the frequency of SV in the sixteenth and seventeenth centuries, it may not be sufficient to simply compare the rates of SV and VS, that is, the possible positions for LEXICAL subjects. In a language like German, any fronting of a nonsubject constituent will result in VS (XVS). In a null-subject language, however, when a nonsubject constituent is moved to preverbal position, the subject, of course, will not necessarily remain postverbal; it may be null.

We suggest that the examination of XV in general (i.e. XVS plus XV-pro) permits an alternative way of detecting V2 in Portuguese. This has been done by Cavalcante and colleagues (2010), as can be seen in Figure 6, where V2 sentences in which the preverbal phrase is not a subject are separate from V2 sentences that are superficially SV. In the graph, 'V2 = XV(S)' stands for any nonsubject constituent preceding the verb, be it the

---

[15] This was first argued by Antonelli (2011). For further arguments, see Galves & Gibrail 2018.

[16] It is not the purpose of this article to fully assume a theory of the syntax/discourse interface. The only crucial assumption is that some movement processes are discursively motivated, which is consistent with both cartographic models (Rizzi 1997) and mapping models (Büring 2007). We argue that the rule that displaces phrases to the preverbal position in ClP is discursively motivated, which entails that it does not apply when the relevant discursive feature/interpretation is not present. See n. 22 for more discussion.

[17] The same fact is shown by Sitaridou (2012) for Old French, Old Portuguese, and Old Spanish.

object or other, and be it with a lexical postverbal subject or a null subject; and 'V2 = SV(X)' stands for superficial V2 order where a subject precedes the verb.
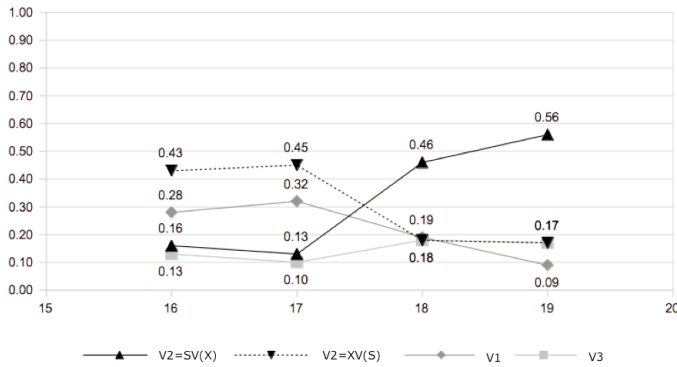

FIGURE 6. Evolution of V1, V2 (SV and non-SV), and V3 in the diachrony of Portuguese, by century
(adapted from from Cavalcante et al. 2010).

As Fig. 6 shows, in the sixteenth and seventeenth centuries, the rates of XV are much higher than the rates of SV—43% and 45%, respectively, as opposed to 16% and 13%. In the eighteenth and nineteenth centuries, while the percentage of SV rises to 46% and 56%, respectively, the XV orders (where X is not a subject) drop to 18% and 17%. To sum up: in V2 orders in general, ClP displays more sentences in which the preverbal phrase is not a subject than sentences in which the preverbal phrase is a subject. It is worth emphasizing that this proportion (43–45% nonsubject XVs) is higher than what has been reported for German (30% nonsubject XVs). If we look at individual authors, the XV data can also be very interesting, as sixteenth- to seventeenth-century authors with high rates of SV have even higher rates of XV: Pinto, for example, presents 51% XV order (against 20% SV, 28% V3, and 1% V1), according to Cavalcante and colleagues (2010).

This means that, even taking into account non-V2 orders—V1 and V3—and even taking into account the possibility of null subjects, the data for ClP shows a definite preference for features that characterize V2 languages: the (X)VS order (which is quite high: 21–35%), and the XV order. In fact, we consider the evidence for verb movement to C(omp) in this data to be very robust.

There is a final crucial point: if our hypothesis that ClP is a grammar in which preverbal phrases are fronted only for discursive reasons is correct, we expect a high amount of variation across authors. This is exactly what we find, as we have shown in §2 (Table 1 and Fig. 3), when we pointed out that the (X)VS/SV variation among sixteenth- to seventeenth-century authors is considerably wider than the variation among modern authors.

The natural question then is whether there is a relationship between the null-subject property and the lack of formal movement. Antonelli (2011) argues that in ClP, the EPP features of C are satisfied by the phi-features on the verb, in the same way the EPP features of T are in non-V-to-C pro-drop languages. From this point of view, if formal movement derives from, or is nothing but, EPP, as claimed by both Frey (2006) and Light (2012), the absence of formal movement in ClP is a consequence of pro-drop.[18]

---

[18] Alternatively, one could argue that it is satisfied by the phi-features of the verb. Thanks to Anthony Kroch for drawing our attention to this point.

We can now go back to the issue of the instability of V2 in pro-drop languages, and more specifically to what caused its loss in the space of a generation in ClP.

3.3. The change from classical to modern european portuguese. As initially observed by Paixão de Sousa (2004), the superseding of postverbal subjects by preverbal subjects is registered in Portuguese texts at the same time as the beginning of the change in clitic placement, which led to the intricate modern pattern. In effect, European Portuguese differs from other Romance languages in that in nondependent, affirmative tensed clauses where the verb is not preceded by a focalized or interrogative phrase, clitic pronouns are categorically enclitic (see Galves & Sandalo 2012 and references therein).[19] We illustrate this fact with preverbal subjects, since this case is highly relevant in this article.

(11) a. O   Paulo **falou=me**.
          the Paulo speak.3sg.pst=1sg.dat.cl
     b. *O   Paulo **me**   **falou**.
          the Paulo 1sg.dat.cl speak.3sg.pst
          'Paulo spoke to me.'

Until the end of the nineteenth century, however, there was variation between proclisis and enclisis in the currently obligatory enclitic contexts, as shown in Figure 7.[20] In the contexts where ClP presents variation, proclisis was the preferred pattern in sixteenth- and seventeenth-century texts (cf. Paixão de Sousa 2004, Galves et al. 2005, and Fig. 7 below). From 1700 on, enclisis with preverbal subjects increases sharply, as Fig. 7 shows for the corpus of the present study, confirming the general picture that had been produced by Paixão de Sousa (2004) and Galves and colleagues (2005).
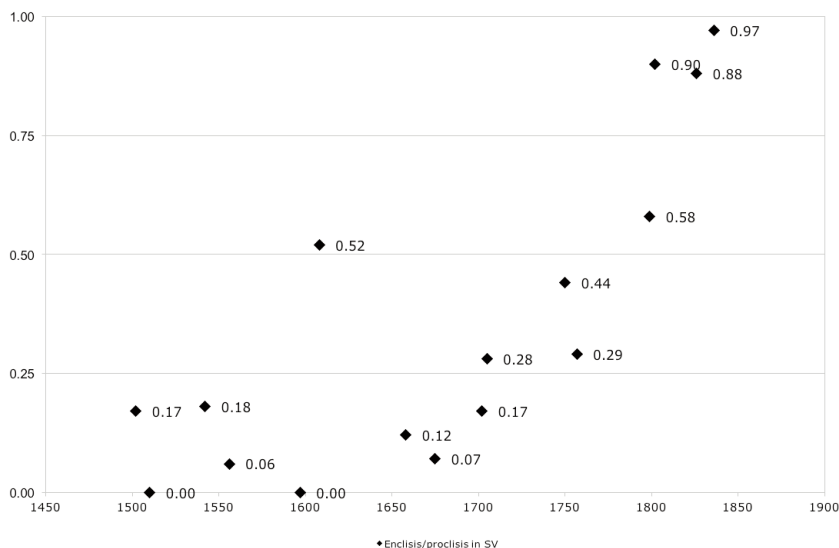

Figure 7. Evolution of the rate of enclisis in clauses with preverbal subjects (SV-cl/S-cl-V).

---

[19] Two comments are relevant here. First, there are no subject clitics in Portuguese. Clitic pronouns correspond only to accusative and dative complements. Second, in the non-V1 contexts in which there is no variation—that is, embedded clauses and matrix clauses in which the verb is preceded by a focalized or interrogated phrase, by some adverb(s), or by negation or a negative phrase—proclisis has been obligatory for the entire attested history of European Portuguese. This means that the obligatorily proclitic placement is not affected by the change that we are considering, which concerns only the contexts in which there was variation in ClP.

[20] Except when the verb was in absolute first position, which is a context in which enclisis has been obligatory for the whole history of European Portuguese.

A good summary of this data would be one that states that texts written by authors born after 1700 reveal enclisis rates of 17–97% in SV sentences, while texts written by authors born before 1700 exhibit enclisis rates of 0–18% in this context. This would be so, if it were not for one seventeenth-century text: Antonio Vieira's *Sermões* 'Sermons', with a 52% enclisis rate in SV, a rate that would seem to put him closer to modern writers than to his contemporaries.[21] It is at this point that the quantitative analysis has to be complemented by closer scrutiny of the data in context. As shown in Galves 2002 and Galves et al. 2005, Vieira's high rate of enclisis is correlated with a stylistic option. We summarize this analysis here.

Galves 2002 and Galves et al. 2005 showed that SV with enclisis in Vieira's sermons corresponds to sentences in which the preverbal subject strongly contrasts with another term in the sentence. This can be seen in the examples below, in which *elles* 'they' and *Christo* 'Christ', and *Deus* 'God' and *os homens* 'the men', respectively, are contrasted.

(12) Elles **conheciam**=**se**,   como homens,
     they know.3PL.PST=3PL.REFL.CL as   men
         Christo **conhecia**=**os**,   como Deus.
         Christ   know.3SG.PST=3PL.ACC.CL as   God
     'They knew themselves as men, Christ knew them as God.'   (Vieira, 1608)

(13) Deus **julga**=**nos**   a nós por nós;
     God judge.3SG.PRS=1PL.ACC.CL to us   for us
         os homens **julgam**=**nos**   a nós por si.
         men   judge.3PL.PRS=1PL.ACC.CL to us   for themselves
     'God judges us for ourselves; men judge us for themselves.'   (Vieira, 1608)

As shown in those same studies, Vieira's sermons are a clear-cut case, because all of the SV-cl constructions are very clearly interpretable as contrastive topicalizations. As is also discussed there, Vieira's patterns of clitic placement in the sermons contrast sharply not only with the patterns in the texts written by his contemporaries, but also with his own patterns in another text—his letters, in which he presents 100% proclisis in SV. Finally, these studies pointed out that in the sermons, proclisis is used not only when the preverbal phrase is focalized, as claimed by Martins (1994), but also in sentences in which the preverbal topic is not markedly contrastive, as in the examples below.

(14) **O Evangelho o**   **diz**:   Erunt signa in sole, et luna, …
     the Gospel   3SG.ACC.CL say.3SG.PRS
         'The Gospel says it: Erunt signa in sole, et luna, … '

(15) **Estes tesouros**, pois, que agora estão cerrados, **se**   **abrirão**
     these treasures thus that now are   enclosed 3PL.REFL.CL open.3PL.FUT
         a seu tempo
         at their time
         'These treasures, thus, that are now enclosed, will open at their own time'

The preverbal phrases in these examples may be termed, following Frascarelli and Hinterhölzl (2007), 'familiar' topics in 14 and 'continuing' topics in 15, both distinct from the contrastive topics in 12–13.[22] In previous works (Paixão de Sousa 2004, Galves et

al. 2005), it was argued that in ClP, the alternation between enclisis and proclisis derives from two different positions of the preverbal phrase. In one of them, the preverbal phrase is 'external' to the clause boundary, and the verb is in a structural first position. In the other, the preverbal phrase is 'internal' to the clause. In both cases, the verb is in a high position, arguably C, as in V2 languages. This is represented in 16, from Galves et al. 2005:52.

(16) a. X [V] → XVcl
b. [X V] → XclV

Following Galves and Sandalo (2012), we now suggest that the boundary that is relevant for clitic placement in ClP is not syntactic but prosodic and that the alternation between enclisis and proclisis in this language is a case of the application of the Tobler-Mussafia law, which prevents clitics from appearing in the first position of the intonational phrase. From this point of view, and adopting a cartographic approach to the left periphery, the difference in the position of the preverbal phrases can be derived from the different positions made available by the different topic nodes. Frascarelli and Hinterhölzl (2007), for instance, argue that the topics that occupy the higher positions of the left periphery of sentences, namely aboutness topics and contrastive topics, are associated with prosodic contours that are independent from the rest of the clause. This nicely accounts for the effect on clitic placement observed in Vieira's sermons.

Figure 8 shows that the decrease of (X)VS and the increase of SV-cl go hand in hand in the corpus analyzed in the present article as well.
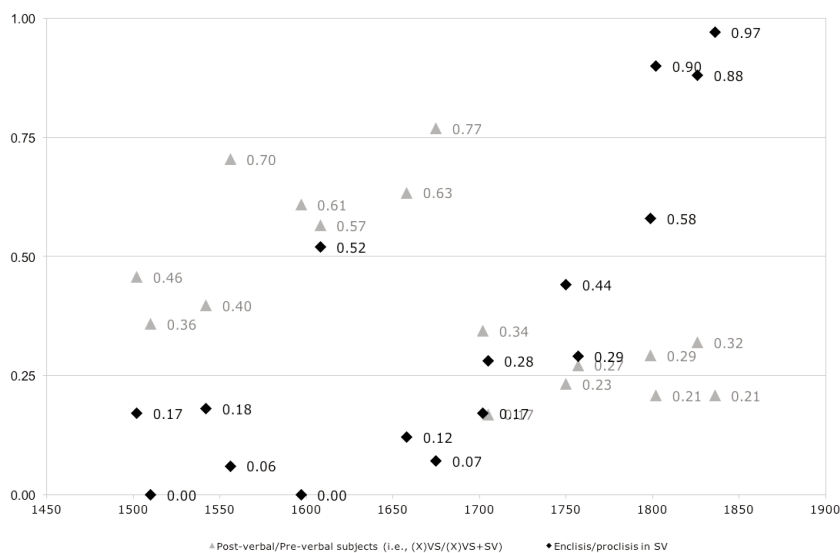

FIGURE 8. Comparative evolution of the rate of postverbal/preverbal subjects and
the rate of enclisis in clauses with preverbal subjects.

responsible for TAB in Germanic languages is *kontrast*, defined by Vallduví and Vikuna (1998:83) as follows: 'If an expression *a* is kontrastive, a *membership set M* = {…, *a*, …} is generated and becomes available to semantic computation as some sort of quantificational domain'. Note that this definition fits nicely into the analyses proposed above for both the proclitic sentences 5 and 6 on the one side, and the enclitic sentences 12 and 13 on the other. In both cases, in effect, the subjects are interpreted with reference to other possible actors involved in the event. This is not true for postverbal subjects. From this point of view, the difference in the interpretation of preverbal phrases with enclitics and proclitics would then have to derive from an additional feature associated with kontrast—perhaps exhaustivity. We leave this complex issue for future research.

Until 1700, SV-cl clauses are very marked constructions. We have seen above that, with the exception of Vieira's sermons,[23] enclisis ranges from 0% to 18% in matrix SV sentences in the authors from the sixteenth and seventeenth centuries. In the authors born in the first half of the eighteenth century, we see that enclisis with preverbal subjects occurs at frequencies of 17% (Cavaleiro, b. 1702), 28% (Aires, b. 1705), and 44% (Marquesa de Alorna, b. 1750). After that, the frequency of SV-cl reaches almost 100% in the last author of our corpus, Ramalho Ortigão, born in 1836. The hypothesis we put forth here to explain this is that the increase in enclisis with preverbal subjects (as well as with other preverbal phrases, although at a lower rate; see Galves et al. 2005) is the effect of a phonological change that affected European Portuguese at some point in the seventeenth century (cf. Frota et al. 2012 for an analysis based on the Tycho Brahe corpus)[24] and favored enclitic placement. In the following sections, we provide more details on this change and propose a model of its effect on clitic placement, and then on the reinterpretation of the position of the subject.

THE PROSODIC CHANGE. The phonological change that is at the source of the current Portuguese prosodic pattern was initially noted due to its impact on the pronunciation of pretonic vowels. This effect was registered for the first time in a *Grammaire Portugaise* published in Paris in 1682 (cf. Teyssier 1980), which mentioned that words like *cortar* 'to cut' were sometimes pronounced as /curtar/, with the raising of the pretonic vowel /o/ to /u/. This was a novelty because until then, this kind of raising was provoked by vocalic harmony, which is clearly not the case in /curtar/ since the last vowel is open. Teyssier (1980) also mentions a *Compendio de orthographia* published in 1767, which contains a list of errors like 'murar' instead of *morar* 'to live', 'purtagem' instead of *portagem* 'toll', 'tucar' instead of *tocar* 'to touch'. Marquilhas (2000:260), studying texts written during the seventeenth century by noneducated people, argues that the frequent use of epenthetic [e] in words like *estragado* 'spoiled', written as 'es-t*e*ragado', can be a reflex of the fact that a word like *parecer* 'to seem' is already pronounced as /parcer/, that is, with a very reduced expression of [e], close to its total deletion, like in EP. It is in the nineteenth century that the extreme realization of the phenomenon of vowel reduction, that is, vowel deletion, is observed by the phonetician Gonçalves Vianna (apud Révah 1956), who complains that the actors of his time are no longer able to properly read the verses of the sixteenth-century poet Camões with all of their syllables, as illustrated in 17a,b.

(17) a. E   se vires   que pode merecer=te.
     b. E   se vir's   que pode mer'cer=te.
          and if see.2SG.FUT that can   deserve.INF=2SG.ACC.CL
            'And if you see it may deserve you.'

According to Gonçalves Vianna, a nineteenth-century Portuguese speaker pronounces the original 17a like 17b, where two /e/s are deleted, and, consequently, two feet are missing. He clearly understands this change as affecting the rhythm of the language, when he comments on the derived difference between European Portuguese and other Romance languages in the following terms:

---

[23] On the exceptional behavior of clitic placement in Vieira's sermons, see also Galves 2002, Galves et al. 2005, and Galves & Sandalo 2012.

[24] Frota and colleagues (2012) look at the distribution of the words in seventeen texts from the Tycho Brahe corpus, according to their size and stress pattern. They find a significant increase in the frequency of monosyllables and oxyton words in authors born from the seventeenth century on. They interpret this fact as the integration of new rhythmic properties in the language, characteristic of stress-timed languages.

> The distance between stressed and unstressed vowels, that is, the difference of intensity between them, may be big or small. Thus, the difference between stressed and unstressed in Germanic languages is maximal; smaller is the difference in Portuguese; even smaller is the difference in Spanish; and minimal is the difference in French. (Gonçalves Vianna 1892:16, apud Frota et al. 2012)

Vianna's comment is very much reminiscent of the distinction between stress-timed and syllable-timed languages, and expresses the idea that EP moved away from the latter and came closer to the former. As argued by Revah (1956), 'this is the most serious modification of the Portuguese pronunciation after the sixteenth century because it modified the very structure of words'. Another related property that can be inferred to have emerged in this change is the tendency to align stress and word boundary both in production and in perception. Several works have shown that, in contrast with what happens in Brazilian Portuguese—which was not affected by the prosodic change—secondary stresses in EP tend to fall on the first syllable of the word (cf. Frota & Vigário 2001 and Sandalo et al. 2006 for a computational implementation based on optimality theory). The crucial point for the present discussion is that this obviously conflicts with the preverbal placement of clitics since, being unstressed by nature, they are not good candidates to receive secondary stress.

   The precise location of this phonological change in time is important if we want to sustain the hypothesis that it triggered the syntactic change that can be seen in the authors born at the beginning of the eighteenth century. Beyond the scarce testimonies by grammarians, and the complex interpretation of facts such as those commented on above, this information is not easy to provide. Frota and colleagues (2012) tried to detect the moment of the change using aspects of prosody easily retrievable from written texts: the syllabic and stress structure of the words. Applying a Bayesian approach to the evolution of the data, they concluded that the pattern observed in the seventeenth century was significantly different from that of the sixteenth century and identical to the eighteenth- and nineteenth-century patterns. This means that a change had occurred at some point in the seventeenth century. This is a welcome result since it is consistent with the initial testimonies about the segmental correlates of the change.

   A MODEL OF PROSODY-DRIVEN SYNTACTIC CHANGE. The analysis of the phonological change affecting pretonic vowels proposed above leads us to hypothesize that the frequency of enclisis in variation contexts grew significantly in the second half of the seventeenth century in spoken language, although this did not appear yet in written texts.[25] The second crucial step is that the increase of enclitic placement has the effect of making the Subject V-cl order unmarked. Hinterhölzl (2009:50–51) claims that:

> syntactic structures are not marked per se (say, in terms of complexity), but count as marked or unmarked if they realize marked or unmarked prosodic patterns. Since the unmarked word order in a language is defined by the predominant, that is to say, the most frequent prosodic pattern in a language, a change in frequency of use of a prosodic pattern can lead to a change in unmarked word order.

If he is right, this means that XP V-cl—and crucially, SV-cl—ceases to be a marked word order at some point in the diachrony of European Portuguese. As we see below, this implies that the initial XP is no longer obligatorily associated with an independent intonational contour. This has two consequences: (i) the reanalysis of the preverbal subject as occupying a subject position and no longer a contrastive topic position, and (ii)

---

[25] Notice, however, that our corpus as it stands is too limited to allow us to draw reliable conclusions relative to the second half of the seventeenth century. Paixão de Sousa (2004) notes that the texts studied by Marquilhas (2000) behave like the other texts from the Tycho Brahe corpus as far as subject position is concerned, but are more enclitic.

the change in the restriction on clitic-first. If the subject and the verb are in the same in-tonational phrase, the restriction on clitic-first can no longer be derived from the Tobler-Mussafia law.[26] Summing up, the model can be outlined in the following way.[27]

(i) Enclisis in variation contexts is favored by the new prosodic pattern →
(ii) Increase in frequency of SV-cl →
(iii) Loss of markedness of SV-cl →
(iv) Reanalysis of the position of the subject in SV-cl →
(v) Loss of V-to-C and reanalysis of enclisis.

In the following sections, we give further empirical evidence for this scenario.

FURTHER EVIDENCE OF THE LOSS OF MARKEDNESS OF ENCLITIC CONSTRUCTIONS. As previously mentioned, Cavaleiro (b. 1702) and Aires (b. 1705) are likely to represent the beginning of a curve of change. However, this is not evident from the purely quan-titative data—their rates of enclisis (17% and 28%, respectively) are not particularly higher than some authors from the previous generations. Their rates are in fact lower than Vieira's in the seventeenth century (52%). Again, a closer look at the data is neces-sary. The comparison with Vieira's sermons is very instructive in this respect, because, as we saw above, all cases of enclisis in the sermons correspond to sentences in which the preverbal phrase is interpreted as a contrastive topic. When the preverbal phrase is a noncontrastive topic, however, proclisis appears without exception. In Cavaleiro, by contrast, there is at least one case in which the preverbal subject of an enclitic sentence is clearly not contrastive.

(18) **Estes meios empregavam=se**   de duas sortes e   tinham dois
     those means use.3PL.PST=3PL.REFL.CL of two kinds and had   two
         nomes.
         names
     'Those means were used in two ways and had two names.' (Cavaleiro, 1702)

In 18, *Estes meios* 'these means' is a continuing topic, which shows that preverbal sub-jects with enclisis can be assigned a neutral interpretation. An analysis of the corre-spondence of Marquesa de Alorna (b. 1750) shows that this is a tendency for the generations to follow. In 19–20 we see that both proclisis and enclisis can be used with a subject whose interpretation is that of a continuing topic.

(19) **Esta reflexão lhe   basta**   para saber   quem
     this   reflection 3SG.DAT.CL suffice.3SG.PRS for   know.INF who
         deve   aceitar   ou recusar
         must.3SG.PRS accept.INF or refuse.INF
     'This reflection suffices in order for you to know who to accept or to re-
         fuse'   (de Alorna, 1750)

---

[26] We refer to Galves & Sandalo 2012 for a proposal on the nature of the new restriction.

[27] This model can be understood as a specific instantiation of the acquisition-based model proposed by Roberts (2007:226–35), where steps (i)–(iii) correspond to E-language changes and steps (iv)–(v) to the new I-language emerging on the basis of the new E-language in the subsequent generations. From this point of view, although the phonology-syntax interface we assume is the one defined in the classical generative model, with phonology interpreting syntax, we believe that phonology can affect syntax inasmuch as it is able to af-fect the E-language serving as primary linguistic data in acquisition. We also assume the view, argued for by many scholars in the past few decades, that suprasegmental phenomena like intonation and rhythm play a cru-cial role in guiding children to their grammar (cf. Morgan 1986, Mehler & Nespor 2002).

(20) **Esta resposta aclarou=me** …
　　　this　answer　clarify.3SG.PST=1SG.ACC.CL
　　　　'This answer enlightened me … '   (de Alorna, 1750)

The variation in 19–20 illustrates the competition between ClP and EP. Example 19 is representative of the former, with the preverbal phrase in a topic position prosodically integrated with the rest of the clause, and the clitic, consequently, proclitic. Example 20, by contrast, is produced by EP, with the preverbal phrase in subject position and obligatory enclisis.

　　An additional novelty shows up in the texts of the first generation of the eighteenth century with regard to the use of enclisis in V3 sentences. Galves and Paixão de Sousa (2005) observe, with a partially different corpus, that in the sixteenth and seventeenth centuries enclisis is possible with V3, and like with V2, it is marginal. However, out of the three possible orders—XXV, SXV, and XSV (X = nonsubject)—the last one, with the subject between a nonsubject and the verb, is never found with enclisis,[28] although it is frequent with proclisis. This changes with the first authors of the eighteenth century, in which we do find the XSV-cl order, as illustrated in 21 and 22. This fact suggests that a new position for the subject was made available, which is no longer sensitive to the restriction observed before.

(21) e　verdadeiramente **estes dois homens sós　divertem=me**
　　　and truly　　　　　these two　men　　alone amuse.3PL.PRS=1SG.ACC.CL
　　　　mais que　a　comédia toda junta.
　　　　more than the comedy　all　together
　　　　'and truly these two men alone amuse me more than all the comedy.'
　　　　　　　　　　　　　　　　　　　　　　　　　(Cavaleiro, 1702)

(22) por isso **os　nossos afectos mudam=se**
　　　for this the our　　affects　change.3PL.PRS=3PL.REFL.CL
　　　　'for this reason our sentiments change'   (Aires, 1705)

　　In sum, we take the two novelties described so far (namely the emergence of XSV-cl and the use of SV-cl where S is not a contrastive topic) as evidence that a new SVO

---

[28] It must be noted that Galves and Paixão de Sousa (2005) found one case of XSV-cl, in Bernardes, born in 1644.

　　(i) Vendo tão rara e verdadeira amizade,
　　　　'Seeing such rare and true friendship,'
　　　el-rei Dionísio o mais velho disse-lhes: …
　　　　'the king Dionísio the older said-to-them: … '

After the article was published, the authors discovered that this sentence had not been appropriately analyzed, due to a modification of the punctuation performed by a twentieth-century editor of the version used in the Tycho Brahe corpus. The punctuation of the original edition is as follows.

　　(ii) Vendo tão rara, e verdadeira amizade
　　　　'Seeing such rare, and true friendship'
　　　el-rei Dionísio o mais velho, disse-lhes: …
　　　　'the king Dionísio the older, said-to-them: … '

As can be seen, in the twentieth-century edition there is a comma between *Vendo tão rara e verdadeira amizade* and *el-rei Dionísio o mais velho*; this made us interpret this mark as the boundary between the two clauses, and *el-rei Dionísio o mais velho* as the subject of the second clause. With the original punctuation, the phrase *el-rei Dionísio o mais velho* is not separate from *Vendo tão rara, e verdadeira amizade*, but rather from *disse-lhe* … . The comma, as we see it now, indicates that the boundary between the two clauses is at this point. This suggests that there is another analysis for this sentence, in which the DP *el-rei Dionísio* … is not the preverbal subject of the tensed verb in the second clause, but the postverbal subject of the gerundive verb in the first clause (and the subject of the second clause would be null). If this analysis is right, there are no occurrences of XSV-cl order in the corpus.

grammar is in competition with the old V2/V-to-C grammar in the texts of the genera-
tions born after 1700. With regard to clitic placement, the new word-order patterns
evolved gradually—the modern categorical use of enclisis in SV was almost attained in
the nineteenth-century texts. With regard to VS, the modern pattern was already instan-
tiated in the eighteenth century. As we saw in §1, there was a steep decline in the rates
of (X)VS in the first generation of authors born after 1700. The range of around 10%
postverbal subjects in those texts corresponds to postpositions that are produced by cur-
rent EP, and which, according to the literature (Ambar 1992, Costa 2004, among oth-
ers), correspond either to unaccusatives or to the interpretation of the subject as focus
and 'residual' VS (such as inversions with affective operators).[29]

EVIDENCE THAT ENCLISIS IS NO LONGER DERIVED FROM THE TOBLER-MUSSAFIA LAW.
One of the contexts in which there is alternation between enclisis and proclisis in ClP is
when a dependent clause immediately precedes the verb, as illustrated in 23 and 24
from Barros (b. 1675).

(23) Para os   começar a render, **amimou=os**   com
     to   3PL.ACC.CL begin.INF to tame.INF please.3SG.PST=3PL.ACC.CL with
        donativos.
        gifts
     'To begin to tame them, he pleased them with gifts.'

(24) Vendo=o   um Cónego no   adro   daquela antiga Sé,
     seeing=3SG.ACC.CL a canon in.the square of.that ancient church
        lhe   disse
        3SG.DAT.CL say.3SG.PST
     'As a canon saw him in the square of that ancient church, he told him: … '

In ClP, the variation between proclisis and enclisis in this context is different from what
we observed above with other preverbal constituents. It is subject to great variation from
author to author, and some of them are highly enclitic in this case (cf. Galves et al. 2005).
From 1700 on, it follows the general tendency toward the generalization of enclisis, and
like in the other contexts, the variation continues until the nineteenth century. The im-
portant point for our discussion is that in the texts written by authors born in the sixteenth
and seventeenth centuries, the placement of clitics is sensitive to the length of the pre-
verbal clause, while it is no longer so in the authors born from 1700 on, as shown by Table
2 (adapted from Galves & Kroch 2016). This is exactly what we predict if the Tobler-
Mussafia law is active in the first period and no longer plays a role in the second one.

	BEFORE 1700		AFTER 1700	
	%	(N)	%	(N)
Clauses ≤ 8 words	47%	(209)	68%	(148)
Clauses > 8 words	60%	(70)	71%	(59)
% enclisis long − % enclisis short	13%		3%	

TABLE 2. Effect of preposed clause length on frequency (%) of enclisis by period.
(From Galves & Kroch 2016. © 2016 John Wiley & Sons, Inc.
Used with permission.)

Table 2 shows that, before 1700, enclisis occurs at a frequency of 60% in clauses con-
taining more than eight words, while in clauses containing eight words or fewer, it occurs
at a rate of 47%. The difference between the two contexts is therefore 0.13. By contrast,

[29] Eide (2010:150) shows that there is a difference in the decrease in the rate of VS according to whether S
is new or old information: VS where S is new information remains at higher rates until the end of the corpus
in this study, which extends until the end of the twentieth century.

after 1700, this difference is close to zero. This reinforces the claim that what governs enclisis from then on is no longer a phonological restriction dependent on the position of intonational boundaries and that XV-cl has ceased to be a marked construction.[30]

4. Concluding remarks. In this article, we have proposed an alternative view of the loss of V2/V-to-C in ClP, which derives it from a change in the prosodic basis of the information structure associated with SV-cl sentences. In this respect, we have challenged Yang's claim that V2 is unstable in Romance languages because of the existence of null subjects. The history of European Portuguese shows that this cannot be the case, since null subjects stably coexisted with properties typical of V2 languages (VS and XV) over centuries.

More generally, we argue that what has been dubbed 'instability' of V2/V-to-C plus pro-drop in the literature derives from the fact that pro-drop V-to-C languages lack formal movement, which means that constituent fronting is dependent only on discursive conditions. Since these conditions are in turn linked to prosodic structure, a prosodic change can have a dramatic effect on the perception learners have of the syntactic structure. In this sense, our analysis is close to the pioneering work on Old French by Marianne Adams (1987), who argued that the loss of V2 in this language was due to the loss of the Germanic stress.[31] Here, we argued that the history of European Portuguese provides us with one more case study illustrating prosody-driven change.[32]

It must be noted that the approach defended here allows us to somehow reconcile the two antagonistic lines of analysis concerning the V2/V-to-C phenomenon in Old Romance languages.[33] On the one hand, many authors argue that it is similar in nature to Germanic V2, in the sense that the verb moves to C or one of the categories of the C layer. On the other hand, other authors rely on the frequency of V1 and V3 in those languages to deny their similarity to Germanic languages. Adopting the latter view, Sitaridou (2012: 556) claims that 'V2 word order obtained in Old Romance is an epiphenomenon of information structure'. In the analysis proposed here, we emphasize that V2 order in ClP depends exclusively on discursive requirements, in contrast with Germanic languages, in which V2 is in most cases the result of a purely formal movement. This claim is entirely compatible with Sitaridou's claim above. From our point of view, however, this does not entail that the structure of main clauses in a language like ClP does not share the crucial property of V2 languages, which is THE ACTIVATION OF THE CP LAYER IN AFFIRMATIVE CLAUSES. We have argued that the frequency and the interpretation of VS as well as the rate of XV empirically support the conclusion that it does.

The consequence of this analysis amounts to shifting the difference between Germanic and V-to-C Romance languages from STRUCTURAL versus LINEAR V2 to FORMAL-PLUS-DISCURSIVE versus DISCURSIVE-ONLY MOVEMENT TO SPEC/CP. This accounts for the frequency of V1 in Romance languages, since a phrase only moves to the pre-

---

[30] A referee claims that this analysis derives a categorical fact, clitic placement, from a noncategorical fact, prosodic phrasing, which is odd. But in fact, we do mean that in ClP, clitic placement in nonobligatorily proclitic V2 contexts is as noncategorical as prosodic phrasing, since it depends on it. We also argue, against the analysis put forth by Pilar Barbosa in several articles (see for instance Barbosa 2000), that this ceases to be the case in EP, where, in fact, clitic placement is categorically enclitic in the contexts discussed (cf. n. 35).

[31] Prosody also plays a crucial role in the analysis of the loss of V2 in French by Kroch (2001:712), where the impact of the prosodic change on acquisition is stressed in the following terms: 'It is not clear what caused the change in French prosody but it is clear what effect it had on the evidence for V2 available to learners'.

[32] Following Galves and Galves (1995) and Galves and colleagues (2005), Eide (2010) also assigns prosody a crucial role in the loss of V2 in Portuguese.

[33] Cf. Ribeiro 1995 versus Kaiser 1999, Fieis 2002, Rinke 2009 for Old Portuguese; and Torres Moraes 1995, Galves 1996, Paixão de Sousa 2004, Galves et al. 2005, Galves & Paixão de Sousa 2005, Gibrail 2010, Antonelli 2011 versus Eide 2006 for ClP.

verbal position when there is a discursive motivation for it. With respect to V3, an aux-
iliary hypothesis is needed. We tentatively propose that it has to do with different posi-
tions of the verb in the CP layer. Frascarelli and Hinterhölzl (2007) argue that in
Germanic languages, the verb is in Force. Antonelli (2011) proposes that it is in Fin in
ClP. This would allow for more than one phrase to occur in preverbal position in the lat-
ter, in contrast to the former. Additionally, we derive from this discursive-only move-
ment to Spec/CP the historical fact that V-to-C in Romance is more prone to change
than V-to-C in Germanic languages.

Finally, in §3 we have shown that in the texts written by authors born in the eighteenth
century, there is evidence of competition between two grammars: the old grammar of ClP
and the new grammar of EP. There are two alternative ways of understanding this com-
petition. The first approach would be to interpret the variation registered in the texts as
evidence of variation in the data available for acquisition of a certain generation. This is,
for instance, the point of view of Yang (2002). As we discussed in §3.1, Yang argues that
the acquisition process consists of competition between grammars and that, in the pres-
ence of a heterogeneous environment, children can acquire two grammars, each one as-
sociated with a different weight. He also claims that the combination of two grammars
remains stable in the speaker's mind, but is diachronically unstable, which accounts for
grammar change after several generations. The empirical evidence for this analysis
would be the variation found in historical texts (Yang 2002:132–33). This model of lan-
guage change, however, can be questioned from a conceptual point of view. It is far from
being commonly accepted that children, given certain conditions, select two different
grammars at the end of their acquisition process. Lightfoot (1979) and Clark and Roberts
(1993), for instance, argue that general processes of simplicity or transparency lead chil-
dren, when confronted with ambiguous data, to select the grammar that generates the
simpler structures. A second approach to variation found in historical texts is to interpret
it not as the cause, but rather as the consequence of grammatical change. In this view (see
for instance Kroch 2001), the variation in texts is not taken as being produced by the ex-
istence of two different grammars in the speaker's mind as the result of his acquisition
process, but rather as the sociolinguistic effect in written texts of former grammatical
changes.[34] The competition takes place between the unique grammar selected by the
writers in their acquisition processes, and fragments of old grammars, acquired later, due
to a normative pressure to which written texts are particularly sensitive.

Our analysis of the change between ClP and EP relies on the latter view. We have ar-
gued that the drastic change in the position of subjects was the result of the reanalysis of
SV sentences as having an underlying SVO order, in which the subject no longer occu-
pied a topic position. We have also suggested that the change in clitic placement played
a crucial role in this reanalysis. The two changes, then, are correlated—but we do not
claim that they are the effect of the same parametric change. In fact, if they were, we

---

[34] Cf. Kroch (2001:702): 'the best studied cases of long-term syntactic drift are most plausibly cases of
grammar competition (that is syntactic diglossia) in which the competing forms may differ in social register,
with an unreflecting vernacular variant slowly driving a conservative written one out of use' (emphasis
added). Further down (p. 725), the author adds:

> Given the strong possibility that textual data do not give evidence for the process of language change in
> vernacular, there is a real need for the study of syntactic innovations in living languages, using sociolin-
> guistic methods to observe unreflecting speech. Such studies do not at present exist, in part because syn-
> tactic change is relatively rare and hard to catch on the fly. In their absence, we can construct abstract
> models of change … . These are useful hypotheses, no doubt, but unless they can be further specified to
> make empirically testable predictions, they will remain speculative. Finding a way to derive such pre-
> dictions is a major task for the future of diachronic syntax.

would expect their curves of evolution to be constant. According to the CONSTANT RATE HYPOTHESIS (Kroch 1989, 2001:721), 'the rate of change in different surface contexts reflecting a single underlying parameter change is the same'. This is not the case in the change in subject position and the change in clitic placement in Portuguese. As can be seen when taking together Figs. 2, 3, 4, and 7, the rate of change in clitic placement is different from the rate of change in the position of subjects: while the latter happens in the span of a generation, the former takes almost two centuries before reaching completion. In fact, given the data presented in this article, the constant rate hypothesis forces us to conclude that the change in subject position and the change in clitic placement are not the effect of the same parametric change. It goes beyond the limits of this article to discuss in detail the exact nature of the underlying changes involved.[35] What we present here, in this regard, concerns the patterns of the evolution of each phenomenon in the texts. We take the difference between the evolutions of enclisis and SV to be derived from the fact that clitic placement is a highly salient phenomenon, to which speakers are sensitive. Additionally, once the position of preverbal subject changed, the proclitic or enclitic position of the pronoun ceased to be associated with distinct prosodic patterns, and therefore lost any discursive effect on the interpretation of the sentence. In this situation, speakers ceased to discursively discriminate proclisis and enclisis—but, as proclisis was the more frequent choice in the older generations, they continued to use it, and only progressively introduced in the written language the innovative form, which is enclisis. The view of grammar competition in texts as a result of previous grammar changes is therefore crucial to the conclusions we draw from the data on ClP in this article.

REFERENCES

ADAMS, MARIANNE P. 1987. *Old French, null subjects and verb second phenomena*. Los Angeles: University of California, Los Angeles dissertation.
AMBAR, MANUELA. 1992. *Para uma sintaxe da inversão sujeito-verbo*. Lisbon: Edições Colibri.
ANTONELLI, ANDRÉ. 2011. *Sintaxe de posição do verbo e mudança gramatical na história do português europeu*. Campinas: University of Campinas dissertation.
BARBOSA, PILAR. 2000. Clitics: A window into the null subject property. *Portuguese syntax*, ed. by João Costa, 31–93. Oxford: Oxford University Press.
BELLETTI, ADRIANA. 2004. Aspects of the low IP area. *The structure of CP and IP: The cartography of syntactic structures, vol. 2*, ed. by Luigi Rizzi, 16–51. Oxford: Oxford University Press.
BÜRING, DANIEL. 2007. Semantics, intonation and information structure. *The Oxford handbook of linguistic interfaces*, ed. by Gillian Ramchand and Charles Reiss, 445–74. Oxford: Oxford University Press.
CASTRO, IVO. 2006. *Introdução à história do Português*. Lisbon: Edições Colibri.
CAVALCANTE, SÍLVIA; CHARLOTTE GALVES; and MARIA CLARA PAIXÃO DE SOUSA. 2010. Topics, subjects and grammatical change: From Classical to Modern European Portuguese. Paper presented at Grammatical Change and the Expression of Subjects (GCES), University of Regensburg.
CLARK, ROBIN, and IAN ROBERTS. 1993. A computational model of language learnability and language change. *Linguistic Inquiry* 24.299–345. Online: http://www.jstor.org/stable/4178813.

---

[35] We assume that the reanalysis of the position of preverbal subjects involves a shift in the features associated with C and T. As for clitic placement, we adopt Galves and Sandalo's (2012) claim that enclisis is derived by a postsyntactic rule sensitive to the specification of the domain in which clitic pronouns cannot appear in first position. The prosodic change affected the perception of this domain by children, leading to a change in the rule of enclisis, which ceased to be sensitive to intonation (the Tobler-Mussafia law) and became sensitive to syntax, producing the very complex and intricate pattern found in EP. We thus depart from Barbosa's analysis of clitic placement in EP that she argues is dependent on prosody. For a full discussion on this matter, we refer the interested reader to Galves & Sandalo 2012 and Galves & Kroch 2016.

Costa, João. 2004. *Subject positions and interfaces: The case of European Portuguese*. Berlin: Mouton de Gruyter.

Eide, Kristine. 2006. *Word order structures and unaccusative verbs in Classical and Modern Portuguese: The reorganisation of information structure*. Oslo: University of Oslo dissertation.

Eide, Kristine. 2010. Prosody, information structure and word order changes in Portuguese. *Diachronic studies on information structure: Language acquisition and change*, ed. by Gisella Ferraresi and Rosemarie Lühr, 143–60. Berlin: De Gruyter.

Fieis, Alexandra. 2002. (XP) V S em Português Medieval. *Actas do XVII Encontro Nacional da Associação Portuguesa de Linguística, Lisboa, Colibri/APL*, 175–87.

Frascarelli, Mara, and Roland Hinterhölzl. 2007. Types of topics in German and Italian. *On information structure, meaning and form: Generalizations across languages*, ed. by Kerstin Schwabe and Susanne Winkler, 87–116. Amsterdam: John Benjamins.

Frey, Werner. 2006. Contrast and movement to the German prefield. *The architecture of focus* (Studies in generative grammar 82), ed. by Valéria Molnár and Susanne Winkler, 235–64. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110922011.235.

Frota, Sónia; Charlotte Galves; Marina Vigário; Verónica Gonzalez-Lopez; and Bernadete Abaurre. 2012. The phonology of rhythm from Classical to Modern Portuguese. *Journal of Historical Linguistics* 2.173–207. DOI: 10.1075/jhl.2.2.02fro.

Frota, Sónia, and Marina Vigário. 2001. On the correlates of rhythmic distinctions: The European/Brazilian Portuguese case. *Probus* 13.247–75. DOI: 10.1515/prbs.2001.005.

Galves, Antonio, and Charlotte Galves. 1995. A case study of prosody driven language change: From Classical to Modern European Portuguese. São Paulo: University of São Paulo, and Campinas: Universidade Estadual de Campinas, ms. Online: http://www.lllf.uam.es/~clase/acceso_local/prosodydriven.pdf.

Galves, Charlotte. 1996. Clitic-placement and parametric change in Portuguese. *Aspects of Romance linguistics: Selected papers from the 24th LSRL*, ed. by Mario Saltarelli, 227–40. Washington, DC: Georgetown University Press.

Galves, Charlotte. 2002. Syntax and style: Clitic placement in Padre Antonio Vieira. *Santa Barbara Portuguese Studies* 6.387–403.

Galves, Charlotte; Helena Britto; and Maria Clara Paixão de Sousa. 2005. The change in clitic placement from Classical to Modern European Portuguese: Results from the Tycho Brahe corpus. *Journal of Portuguese Linguistics* 4(1).39–68. DOI: 10.5334/jpl.166.

Galves, Charlotte, and Alba Gibrail. 2018. Subject inversion in transitive sentences from Classical to Modern European Portuguese: A corpus-based study. *Word order change*, ed. by Adriana Cardoso and Ana Maria Martins. Oxford: Oxford University Press, to appear.

Galves, Charlotte, and Anthony Kroch. 2016. Main syntactic changes from a principle-and-parameters view. *The handbook of Portuguese linguistics*, ed. by W. Leo Wetzels, João Costa, and Sergio Menuzzi, 487–503. Oxford: Wiley-Blackwell.

Galves, Charlotte, and Maria Clara Paixão de Sousa. 2005. Clitic placement and the position of subjects in the history of European Portuguese. *Romance languages and linguistic theory 2003: Selected papers from 'Going Romance' 2003, Nijmegen, 20–22 November*, ed. by Twan Geerts, Ivo van Ginneken, and Haike Jacobs, 93–107. Amsterdam: John Benjamins.

Galves, Charlotte, and Filomena Sandalo. 2012. From intonational phrase to syntactic phase: The grammaticalization of enclisis in the history of Portuguese. *Lingua* 122.952–74. DOI: 10.1016/j.lingua.2012.03.005.

Gibrail, Alba. 2010. *Contextos de formação de estruturas de tópico e foco no português clássico*. Campinas: Universidade Estadual de Campinas dissertation.

Gonçalves Vianna, A. R. 1892. *Exposição da pronúncia normal portuguesa, para uso dos nacionais e estrangeiros*. Lisbon: Imprensa Nacional.

Hinterhölzl, Roland. 2009. The role of information structure in word order variation and word order change. *Information structure and language change: New approaches to word order variation in Germanic*, ed. by Roland Hinterhölzl and Svetlana Petrov, 45–66. Berlin: Mouton de Gruyter.

Kaiser, Georg. 1999. A ordem das palavras e a posição do verbo finito no português antigo. *Actas do Congresso Internacional por motivo dos vinte anos do português no ensino*

superior, ed. by Ferenc Pál, 248–61. Budapest: Departamento de língua e literatura portuguesas da Faculdade de Letras da Universidade Eötvös Loránd de Budapeste.

Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1.199–244. DOI: 10.1017/S0954394500000168.

Kroch, Anthony. 2001. Syntactic change. *The handbook of contemporary syntactic theory*, ed. by Mark Baltin and Chris Collins, 699–729. Oxford: Blackwell.

Light, Caitlin. 2012. *The syntax and pragmatics of fronting in Germanic*. Philadelphia: University of Pennsylvania dissertation.

Lightfoot, David. 1979. *Principles of diachronic syntax*. Cambridge, MA: MIT Press.

Lightfoot, David. 1997. Catastrophic change and learning theory. *Lingua* 100.171–92. DOI: 10.1016/S0024-3841(93)00030-C.

Marquilhas, Rita. 2000. *A faculdade das letras*. Lisbon: Imprensa Nacional Casa da Moeda.

Martins, Ana Maria. 1994. *A história dos clíticos em português*. Lisbon: University of Lisbon dissertation.

Mehler, Jacques, and Marina Nespor. 2002. Linguistic rhythm and the acquisition of language. *Structure and beyond: The cartography of syntactic structures, vol. 3*, ed. by Adriana Belletti, 213–22. Oxford: Oxford University Press.

Morgan, James. 1986. *From simple input to complex grammars*. Cambridge, MA: MIT Press.

Namiuti, Cristiane. 2008. *Aspectos da história gramatical do português: Interpolação, negação e mudança*. Campinas: University of Campinas dissertation.

Paixão de Sousa, Maria Clara. 2004. *Língua barroca: Sintaxe e história do português nos 1600*. Campinas: University of Campinas dissertation.

Révah, Israel. 1956. L'évolution de la prononciation au Portugal et au Brésil du 16éme siècle à nos jours. *Anais do primeiro congresso brasileiro da língua falada no teatro, Salvador*, 387–402.

Ribeiro, Ilza. 1995. Evidence for a V2 phase in Old Portuguese. *Clause structure and language change*, ed. by Adrian Battye and Ian Roberts, 110–39. Oxford: Oxford University Press.

Rinke, Esther. 2009. Verb-placement in Old Portuguese. *Focus and background in Romance languages*, ed. by Andreas Dufter and Daniel Jacob, 309–32. Amsterdam: John Benjamins.

Rizzi, Luigi. 1997. The fine structure of the left periphery. *Elements of grammar: A handbook of generative syntax*, ed. by Liliane Haegeman, 281–337. Dordrecht: Kluwer.

Roberts, Ian. 2007. *Diachronic syntax*. Oxford: Oxford University Press.

Sandalo, Filomena; Maria Bernadete Abaurre; Arnaldo Mandel; and Charlotte Galves. 2006. Secondary stress in two varieties of Portuguese and the Sotaq optimality based computer program. *Probus* 18(1).97–125. DOI: 10.1515/PROBUS.2006.004.

Sitaridou, Ioanna. 2012. A comparative study of word order in Old Romance. *Folia Linguistica* 46(2).553–604. DOI: 10.1515/flin.2012.019.

Teyssier, Paul. 1980. *Histoire de la langue portugaise*. Paris: Presses universitaires de France.

Torres Moraes, Maria Aparecida. 1995. *Do Português Clássico ao Português Europeu Moderno: Um estudo diacrônico da cliticização e do movimento do verbo*. Campinas: University of Campinas dissertation.

Vallduví, Enric, and Maria Vilkuna. 1998. On rheme and kontrast. *The limits of syntax*, ed. by Peter Culicover and Louise McNally, 79–108. New York: Academic Press.

Yang, Charles. 2000. Internal and external forces in language change. *Language Variation and Change* 12.231–50.

Yang, Charles. 2002. *Knowledge and learning in natural languages*. Oxford: Oxford University Press.

Galves
Bloco IV - Pavilhão Docentes - térreo
Sala D.1.36 - Sala de Docente
Rua Sérgio Buarque de Holanda, no 571
CEP 13083-859 - Campinas - SP, Brazil
[galvesc@unicamp.br]
[mariaclara@usp.br] (Paixão de Sousa)