THE LOSS OF VERB-SECOND IN THE HISTORY OF PORTUGUESE: SUBJECT POSITION, CLITIC PLACEMENT AND PROSODY

CHARLOTTE GALVES
Universidade de Campinas / Instituto de Estudos da Linguagem

MARIA CLARA PAIXÃO DE SOUSA
Universidade de São Paulo / Faculdade de Filosofia, Letras e Ciências Humanas

ABSTRACT. This paper analyses the changes in subject position patterns in the diachrony of Portuguese between the 16th and 19th centuries in terms of a grammatical change involving the loss of a Verb-second system and the raise of an SVO system. This analysis is based on a corpus of syntactically annotated texts written by Portuguese authors born between 1502 and 1836, which represents the largest set of data available for this period of the language. Based on this data, we argue that in Classical Portuguese – i.e., the grammar instantiated in 16th-17th century texts – the pre-verbal position is reserved for discourse prominent constituents, with a strong correlation with prosodic contours, whereas in Modern European Portuguese – i.e., the grammar instantiated in 18th-19th century texts – the pre-verbal position is reserved for subjects. Our analysis for the grammar of Classical Portuguese relies on an alternative account of the “V2” property in null-subject languages, thus potentially bearing on studies about V2 and its loss in Old Romance in general. In the case of Portuguese, we suggest that the change from V2 into SVO is grounded on a prosodic change that took place in the second half of the 17th century, which also affected clitic placement. Following the concept of Grammar Competition in Kroch (2001, a.o.), we conclude by arguing that the evolution of word order patterns in the 18th and 19th centuries reveals the effects of this change gradually becoming visible in written texts.

1. INTRODUCTION

In this paper, we describe the change in subject position and clitic placement in the history of Portuguese between the 16th and 19th centuries, based on data extracted from sixteen...
syntactically annotated texts from the *Tycho Brahe Corpus*, written by Portuguese authors born between 1502 and 1836. This would correspond to what has been called, in the tradition of Portuguese historical linguistics, Middle Portuguese, Classical Portuguese, and the beginning of Modern Portuguese. Here, following the tradition, we call “Classical Portuguese” the language instantiated in texts written by authors born between the 16th and 18th centuries; and our focus is on the change from Classical to Modern European Portuguese. Previous studies have shown that this language corresponds to a grammar that is different both from Old Portuguese and Modern European Portuguese (cf. Galves et al. 2006). Classical Portuguese is a pro-drop, V-to-C language (cf. Torres Moraes, 1995; Galves, 1996; Paixão de Sousa, 2004), with pronominal enclisis in V1 clauses and an enclisis/proclisis alternation in non-dependent V2 clauses (cf. 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 as regards clitic placement (Namiuti, 2008); and Modern European Portuguese is a pro-drop, SVO language, with pronominal enclisis in V1 and obligatory enclisis in SV clauses.

We are therefore offering empirical evidence for the full understanding of the problem of the “instability” of V2 in pro-drop languages, pointed out by Yang (2002). As we will show, the data on Classical Portuguese indicates the need for a review of Yang’s explanation for this problem and its development. Yang (2002) explains the loss of V2 in Romance languages as a consequence of its combination with null subjects, which would make such grammars “intrinsically unstable”. We shall argue, however, that this is not a sufficient account for the loss of V2 in Portuguese, since V2 had coexisted with null subjects for at least five centuries before the emergence of SV. As was argued in previous work (Galves et al. 2005), the loss of V2 in Classical Portuguese is linked to the change in clitic-placement, which in turn is related to a change in the rhythmic pattern of the language (Galves and Galves, 1995). We propose that an increase of enclisis with pre-verbal subjects, related to the new prosodic pattern of the language, contributes to the loss of the markedness previously associated with this order, and to the non-identification of pre-verbal subjects as high topics (as our analysis of the development of XSV-cl will show). This gives a higher weight to the reanalysis of pre-verbal subjects as occupying a pre-verbal subject position and no longer a topic position.

The paper is organized as follows: in Section 2 below, we explore the aspects of word order in Classical Portuguese which we take as essential for the understanding of the grammatical change - subject position and clitic placement. In Section 3, we analyze the dynamics of the change and propose an interpretation for the loss of V2 in Classical Portuguese, and its relation with pro-drop. We argue that, contrary to what happens in Germanic languages, V2 in pro-drop languages is totally dependent on discourse conditions, and consequently on prosody – and therefore, more prone to suffer the effect of prosodic changes. Finally, in the concluding remarks, we point out to some implications of our proposal for the analysis of the difference between Germanic and Romance languages, and for the notion of competition of grammars.

2. Word order in the diachrony of Portuguese

Many of the aspects to be discussed in this section have been tackled to some degree in previous studies - in particular, Paixão de Sousa (2004), Galves & Paixão de Sousa (2005), and Galves et al. (2005), all of which were limited to word order in sentences with clitics. The present paper confirms the results of these previous works, additionally discussing sentences both with and without clitics, amplifying and enlarging the universe of analysis. The present survey was conducted with an unprecedented amount of data for this period of the language, extracted from a syntactically annotated corpus - comprising 34.293 declarative main clauses in texts written by 16 authors born between 1502 and 1836.
As we shall argue further on, we interpret the results of this survey as important evidence of a Verb-second system in the authors of the first phase (i.e., 16th and 17th century authors). However, as with other Old Romance Languages, "V2" in Classical Portuguese 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 and verb-third orders. What we have then, in short, is a Verb-second, null subject grammar in which the verb may surface in first, second or third position. Our central aim in this paper is to model the change between such a system and a canonical SV system – Modern European Portuguese. In accordance with this aim, we will focus on the possible expressions of subjects in Classical Portuguese: lexical (VS or SV) and null; and the varied possible positions for the verb will be shown with reference to the data and discussion presented by Cavalcante et al. (2010), Gibrail (2010) and Antonelli (2011). In addition, given the nature of our main hypothesis - according to which clitic placement plays a key role in the change - we shall discuss the patterns of clitic placement in relation to subject position.

The section begins with an account of the frequency changes documented for lexical and null subjects, proceeds to a description of the main facts as regards clitic placement, and ends with a summary of our general analysis for the two combined sets of evidence.

2.1 Subject Position

The most remarkable fact regarding subject position in Classical Portuguese is the high frequency of VS over SV, as compared do Modern European Portuguese. Figure 1 below shows the rates of VS over SV in 16th-19th century texts, revealing two clear groups: in texts written by authors born in the 16th-17th centuries, VS ranges from 36% to 77%, whereas in texts written by authors born after the 18th century, VS ranges from 17% to 31%:

**Figure 1**: Relative Position of Lexical Subjects in matrix clauses (VS/SV)

We take the high frequency of VS orders in 16th-17th century texts as one of the most important pieces of evidence for a V2 system in Classical Portuguese. The data on subject-postposition depicted in Fig. 1 includes both verb-initial constructions and constructions in
which a pre-verbal element other than the subject appears in the pre-verbal position - XV(S), be it XVS (X fronted, post-verbal Subject), or XV (X fronted, null subject), both quite frequent in these texts. We take both XVS and XV as instances of general "XV", i.e., of "non-SV" orders - in which a constituent other than the subject occupies the pre-verbal position. We claim that in Classical Portuguese, the pre-verbal position is reserved for discursive-prominent elements. We are not concerned, in this paper, with the exact nature of this position (and we refer to Antonelli, 2011, for a concrete proposal in the framework of Rizzi’s cartography).

The relevant point is that Classical Portuguese is not an SV system, and pre-verbal subjects occupy the same position as other fronted elements. In (1-3) below, we illustrate the possible patterns of the expression of the subject, SV (1), VS (2), and null subjects (3):

(1) Christo Senhor nosso, disse a seus Discipulos, que o segredo d’aquelle dia é reservado só ao Padre.

‘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.

‘The king started the church of São Vicente’, (Sousa 1556)

(3) e com a sua prisão mudaram de intento

‘and with his prision, they changed their intent’ (Galhegos, 1597)

The frequency of each of these three possibilities for the expression of subjects in Classical Portuguese is interestingly different from what may be documented for a canonical pro-drop SV system (such as Modern European Portuguese). Figure 2 below 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**: Evolution of the expression of subjects, 16th to 19th centuries (%VS, % SV and % Null S, main clauses) (Grouped by 100 year periods)
As we can see, in 16th-17th century texts the proportion of VS is 21% and 35% respectively, the proportion of SV is 18% and 17%, and null subjects represent 62% and 48% of main clauses. In short, VS in this period is on average higher than SV and lower than nulls. It is only after 1700 that SV becomes a favored order for lexical subjects in Portuguese texts, surpassing VS at rates of 41% against 12% in the 18th century, and 34% against 11% in the 19th. Notice, in particular, that because we are comparing all three subject-placement possibilities for subjects (including null), we can see clearly that the decrease in VS at the turn of the 18th century in fact corresponds to an increase in SV (and not, for example, to an increase in the rate of null subjects, although this rate may oscillate; see below). To sum up:

i. VS decreases between the 17th and the 18th centuries;

ii. SV surpasses VS at this same point;

iii. Null subject proportions vary, but show no clear tendency pattern

Although this general picture very much encompasses the fundamental data to be discussed in section 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-18th century and post-18th century). Table 1 and Figure 3 below show this, detailing the data for SV, VS and null subjects previously shown above, this time for each of the sixteen texts of our corpus:

Table 1: Evolution of the expression of subjects, 16th to 19th centuries

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>%NS</td>
<td>0.66</td>
<td>0.68</td>
<td>0.60</td>
<td>0.68</td>
<td>0.47</td>
<td>0.38</td>
<td>0.60</td>
<td>0.47</td>
<td>0.60</td>
<td>0.53</td>
<td>0.53</td>
<td>0.61</td>
<td>0.53</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>%SV</td>
<td>0.19</td>
<td>0.21</td>
<td>0.24</td>
<td>0.09</td>
<td>0.21</td>
<td>0.27</td>
<td>0.15</td>
<td>0.12</td>
<td>0.29</td>
<td>0.54</td>
<td>0.36</td>
<td>0.34</td>
<td>0.28</td>
<td>0.42</td>
<td>0.36</td>
</tr>
<tr>
<td>%VS</td>
<td>0.16</td>
<td>0.12</td>
<td>0.16</td>
<td>0.22</td>
<td>0.32</td>
<td>0.35</td>
<td>0.25</td>
<td>0.41</td>
<td>0.13</td>
<td>0.11</td>
<td>0.11</td>
<td>0.12</td>
<td>0.11</td>
<td>0.11</td>
<td>0.08</td>
</tr>
</tbody>
</table>
The rates of null subjects vary considerably among contemporary authors in the corpus. We consider this to be expected, as the option for null subjects in a pro-drop language is related to textual, not strictly grammatical constraints (we come back to this point more precisely below). We are more interested here in the variations concerning the position of lexical subjects. Figure 3 shows that 16th-17th century authors present a relatively wide range of variation: VS ranges from 12% to 41%, and SV from 9% to 27%. For 18th-19th century authors a more consistent pattern emerges: VS stays around 10%. In addition, we can see that for all 18th-19th century authors, SV rates are higher than the rates of VS. This is not true for 16th-17th century authors: in this period, some authors present more SV than VS, others, more VS than SV.

So, our first conclusion from this data is that there is more variation in the position of lexical subjects in the first period than in the latter, a fact to which we will come back later. Our second conclusion concerns the patterns of variation within the group of 16th-17th century authors. It is true that we find authors, in the 16th century, that present frequencies of pre-verbal subjects of over 20%, coming close to the 28% threshold of 18th 19th century authors—for instance, Vieira (1608) with 27% SV, Couto (1542) with 24% SV, or even Pinto (1510), with 21% SV. As we will comment later, in spite of that, those texts are clear representatives of Classical Portuguese. We interpret this in the light of the style in which the texts were written. Subsection 2.2 below presents a short discussion on this matter, after we present the data on clitic placement, which, as we have suggested, is crucial to the understanding of the grammatical change into Modern European Portuguese and which was also crucial for our proposal of the nature of verb-second in Classical Portuguese.

As regards the data summarized so far, the important point is as stated above: it is only after the 18th century that we see SV as a predominant pattern in Portuguese texts.
2.2 Clitic Placement

The most important fact on the diachrony of clitic placement from Classical do Modern European Portuguese is that enclisis with subjects increases sharply after the 18th century. Fig. 4 below shows that for the corpus of the present study, confirming the general picture that had been produced by Paixão de Sousa (2004) and Galves et al. (2005):

**Figure 4: Evolution of Enclisis/Proclisis in clauses with pre-verbal subjects (SV-cl / S-clV)**

A good summary of this data would be one that said that texts written by authors born after the 18th century reveal rates of enclisis of 17% to 97% in SV sentences, while texts written by authors born before 1700 exhibit rates of 0% to 18% enclisis in this context. This would be so, if it were not for one 17th century text: Vieira's *Sermons*, with 52% enclisis in SV, a rate that would seem to put him closer to modern writers than to his contemporaries. 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 will summarize this analysis here.

Galves (2002) and Galves et al. (2005) showed that SV with enclisis in Vieira's *Sermons* correspond to sentences in which the preverbal phrase is a contrastive topic. This can be seen in the examples below, in which *Elles* ("them") and *Christo* ("Christ"), and *Deus* ("God") and *os homens* ("the men"), respectively, are contrasted.

(4) Elles *conheciam*-se, como homens,
    They *know.3PP.PAST*=themselves, as men,
    Christo *conhecia*-os, como Deus.
    Christ *know.3PP.PAST*=them, as God

    'They knew themselves as men, Christ knew them, as God.' (Vieira, 1608)

(5) Deus *julga*-nos a nós por nós;
    God *judge.3PS.PRES*=us to us for us;
    os homens *julgam*-nos a nós por si.
    men *judge.3PP.PRES*=us to us for themselves.
As it was shown in those same studies, Vieira's Sermons are a clear-cut case, because the totality of the SV-cl constructions are very clearly interpretable as contrastive topicalizations. As it is also discussed there, Vieira's patterns of clitic placement in the Sermons contrasts 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 has 100% proclisis in SV. Finally, the authors pointed out that in the Sermons, proclisis is used not only when the pre-verbal phrase is focalized, as claimed by Martins (1994), but also in sentences in which the pre-verbal topic is not contrastive, as in the examples below.

    The Gospels say: ... 'The Gospels says it: ...

(7) Estes tesouros, pois, que agora estão cerrados, se abrirão a seu tempo
    These treasures, thus, that now are enclosed, them=will-open.3PP.FUT
    at their time

'These treasures, thus, that are now enclosed, will open at their own time'

In the examples above, the pre-verbal phrases may be termed, after Frascarelli & Hinterhölzl (2007), "familiar topics" in (6) and "continuing" topics in (7), both distinct from the contrastive topics in (4-5). In Galves et al. (2005), it was argued that in Classical Portuguese, contrastive topicalization corresponds to an actual V1 structure, with the topic in an external position, and the verb in first position in the clause; proclisis, on the other hand, would indicate actual V2 structures. Paixão de Sousa (2004), in fact, had argued that XV structures with proclisis represent instances of V2-fronting in Classical Portuguese, where the fronting corresponds to discourse prominence. In both cases, this would include SV, to be interpreted as contrastive topicalization or V2-fronting of subjects, yielding, respectively, Subject-Verb-clitic and Subject-clitic-Verb.

Working in the “Cartographic model” proposed by Rizzi (1997) and subsequent works, Frascarelli & Hinterholzl, (2007) show that the topics that occupy the higher positions of the left-periphery of sentences, namely aboutness topics and contrastive topics, are associated to prosodic contours independent from the rest of the clause. This nicely accounts for the V1 effect observed in Classical Portuguese, and leads to a slightly different interpretation than the one proposed in Galves et al. (2005), in which V1 should be understood as a prosodic notion in accordance to the spirit of the original Tobler-Mussafia law (see Galves and Sandalo 2012 for a detailed analysis in the framework of Distributed Morphology).

The most important point we wish to make here is that, regardless of the exact structural position and of the exact kind of topic, the pre-verbal position in Classical Portuguese is reserved for prominent elements of varied kinds, all of them dependent on the information structure. In the case of written texts (which is what we have in our hands), we may consider the text itself as the discourse context - and that means that different texts will produce different patterns of word orders. Paixão de Sousa (2004) pointed out, for instance, that SV with proclisis is particularly frequent in texts with abundant narrative sequences, in which different participants of the narrated event take their turns in a segment of the text - such as in Vieira's letters, with abundant proclisis, as we mentioned above. This is also the case of the text Peregrinação, by Fernão Mendes Pinto, included in the present study (and with 100% proclisis in SV, cf. Fig. 3). Let us examine the following examples of SV with proclisis in this text:
O capitão-mor lhe respondeu que os embaixadores tinham seguro para suas pessoas.

The captain-major answered him that the ambassadors had personal insurance.

‘The captain-major answered him that the ambassadors had personal insurance’, (Pinto, 1510)

O brâmene lhe deu por isso seus agradecimentos,

The braman thanked him for that.

‘The braman thanked him for that’, (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 braman'. Each time one of the characters is mentioned, so that their turn in the conversation is described, a constituent that occupies the pre-verbal position (i.e., the V2-topic position) is used.

According to our main hypothesis, the structure for SV with or without clitics should follow similar patterns; take the following examples of SV without clitics, from the same text:

(10) nosso capitão-mor cometeu então queimar-lhe a galé,

our captain-major committed then to burn=him the galé,

‘our captain-major proceeded then to burn his galé’, (Pinto, 1510)

In (10), the fronted constituent is also the grammatical subject of the sentence in which it appears; thus, we have the surface order "SV". The same phrasal organization can be seen, naturally, in sequences in which the fronted element is not the subject. In such cases, we will have XV(S) sequences - i.e., XV, with null subjects; or XVS, with post-verbal lexical subjects. Take, for instance, the following sentences from the same text:

(11) Ao mercador que me trouxe mandou Pero de Faria dar sessenta cruzados...

To the merchant that me= bring=3PS-PAST order=3PS-PAST

Pero de Faria to give sixty cruzades...

‘To the merchant who brought me, Pero de Faria ordered that sixty cruzades be given…’, (Pinto, 1510)

Again in this case, we are 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 (11)) take turns in receiving presents and instructions from another, constant character ('Pero de Faria'). Naturally, here, the grammatical subject of the sentences 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 pre-verbal; and the subject, being less prominent, does not.

This is how VS - more specifically, XVS - can be accounted for in Classical Portuguese: instances of V2-fronting of non-subjects. The proportion of SV and VS, therefore, may oscillate considerably as a function of the discourse nature of the texts in which they appear. It will be crucial for us to show that this oscillation conditioned by discourse (and thus, stylistic) requirements is expected to be sharper in a V2 system than in an SV system. As we
showed before in Table 1 and Figure 3, this is precisely what happens - post 18th century texts present less variation among contemporaries than pre-18th century texts.

2.3 Summing up

In broad terms, the analysis proposed for word-order patterns in Classical Portuguese in this paper follows what had been proposed in Paixão de Sousa (2004), Galves & Paixão de Sousa (2005), and Galves, Britto & Paixão de Sousa (2005), to the effect that in this grammar, there are more than one structural position for pre-verbal constituents - with different intonational contours associated, which have effect on the position of the clitics. Critically, neither of these positions is reserved for subjects.

Sentences with clitics are therefore key data for this proposal, as the restriction on clitic-first in the Intonational Phrase correlates S-V-clitic with a marked prosodic pattern. At this point, it will be important to pursue a possible interpretation for the change into Modern European Portuguese allowed by the analysis above. According to this analysis, SV in classical texts (1500-1600) and SV in modern texts (1700-1800) correspond to different structures, and different grammars: in the classical texts, SV correspond to topicalization constructions just like any XV construction; and like any XV, SV will be followed by proclisis or enclisis depending on the nature of the external or internal topic (in other words: here, SV is a subtype of XV, as in any regular V2 grammar). In the modern texts, however, SV corresponds to subjects in canonical positions, which now must be followed by enclisis (due to interface factors with prosody, as we will show below) - in other words: here, an SVO grammar has emerged. Our central question is now: why was the old grammar replaced by the new? In section 3 below, we explore this question.

3. Romance V2 revisited

This section is organized as follows. In 3.1, we present some of the challenges presented by Romance V2 in general to the understanding of language change, and we discuss the solution suggested by Yang (2002). In 3.2, we suggest an alternative approach on Romance V2 in general, based on Frey (2006)’s analysis of German, and in 3.3 we apply it to analyze the change from Classical to Modern European Portuguese in particular.

3.1 Romance V2 and its challenges

Old Romance languages have been often analyzed as verb-second systems with some deviations from the more “canonical”, Germanic V2. First of all, as mentioned by many authors, in Old Romance languages V3 is much more frequent than in Germanic V2 languages (cf. Sitaridou 2012, a.o). In addition, and as we mentioned further above, the coexistence of V2 with null subjects also puts Romance V2 in a different category as compared to Germanic V2, as noted in particular by Yang (2002).

Both properties – null subjects and coexistence with V3– apply to Classical Portuguese, as some of the data discussed in Section 2 has already shown; in 2.1, for instance, we saw that null subjects are prevailing and constant in Classical (and Modern European) Portuguese. Regarding V3, recent research has shown that it is also frequent in Classical Portuguese, and co-exists not only with V2, but also V1 and even some V4. Cavalcante et al. (2010), who work with part of the same corpus as in the present study, showed that V3 sentences occur at the rate of 13% and 10% in the 16th and 17th centuries respectively (as we will comment with more detail further below).

Classical Portuguese, therefore, qualifies as a good case study for Romance V2, and the diachronic data brought by our survey might serve as empirical evidence for some of the theoretical work recently dedicated to solving some of the puzzles presented by such
grammars - in particular Yang (2002), who points out their “inherent instability”. In this section we shall argue, however, that neither null subjects nor the presence of V3 and V1 orders must be regarded as “instabilities” in Classical Portuguese. 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 to distinct prosodic patterns. Speakers of Classical Portuguese could perfectly differentiate 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 shall argue further on, is Prosody.

In order to discuss this, we first present a summarized account of Yang (2002)’s 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, in which when a grammar successfully analyzes an utterance the child is exposed to, it is rewarded; when it fails, it is punished. More formally speaking (Yang, 2002: 26), “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, rewards G_i by increasing p_i; otherwise, punishes G_i by decreasing p_i.”. At each step of the acquisition process, therefore, “each grammar G is paired with a weight p which can be viewed as the measure of prominence of G in the learner's language faculty”. 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. But in case of heterogeneous environments, no grammar is 100% compatible with the data. In this case, the model allows language learners to “form internal representations of coexisting grammars” (Yang, 2002: 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 the effect of general principles of economy that lead children to choose the simplest one (cf. for instance Clark and 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, as time goes by, 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 0. 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 the author, evidence for V2 is VS (XVS0, OVS), and evidence for SVO is V3 (SXVO,XSVO). The SVO order, however, is ambiguous 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 non-ambiguous evidence acquirers of a V2 language have at their disposal has a weight corresponding to this rate. On the other hand, in SVO languages (like English) only 10% of all sentences are V3 patterns – i.e., are non-ambiguous. Following the author’s point of view, this means that the weight of the non-ambiguous evidence for SVO is weaker than the weight of the non-ambiguous 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).” This would account for the strong stability of V2 in Germanic languages, but leave 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 diminished in the total dataset because of the existence of null-subject sentences. The author then observes
that in three texts written in French at the turn of the 15th century, null subjects occur at a rate of c. 40%; as for VS, instead of the 30% found in non-pro-drop V2 languages, he finds a frequency of 5-7-18% depending on the text. In those same texts, V->2 ranges from 11 to 15%. 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 to an SVO (plus pro-drop) grammar” (Yang, 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 in the history of another pro-drop Romance language: European Portuguese.

Our data on Classical Portuguese shown so far evidences an increase of pre-verbal subjects between the 17th and the 18th centuries – as we have shown in Section 2, cf. Fig. 2. This would mark the 1700s as the point in which SVO overcomes V2 in the diachrony of Portuguese. However, this change in the data is not preceded by a situation where VS is overcome by V3: 16th and 17th centuries Classical Portuguese texts present rates of 13% and 10% of V3 sentences, respectively (cf. Cavalcante et al., 2010), whereas VS patterns in those centuries range from 21% to 35% (cf. Fig. 2).

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 overcame V2 patterns even without a change in the relation between V3 and VS – as the rise in SV is not preceded by a significant change in the rates of V3. Always 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.

The fact remains that the patterns of word order present a steep change on the turn of the 1700s, and this drastic change in the space of one generation must be explained. We shall attempt to do this 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 Classical Portuguese and Germanic languages. Such a reflection profits a lot from the analysis proposed by Frey (2004, 2006) for V2 in German and recently discussed and extended to other Germanic languages by Light (2012).

Frey (2004, 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 (EPP) properties of Comp, more exactly Fin, and the other one is topicalization of some element of the clause, what the author calls “True A-bar-movement (TAB)”, to higher positions in the CP field. Frey argues at length that only the A-bar movement entails a pragmatically marked interpretation, which is, according to him, one of contrast. This accounts for the long-noticed fact that the interpretation associated with the pre-verbal position in German is not the same in all cases, which would remain unexplained under a unitary analysis of the movement to Spec/CP. This also accounts for the fact that SVO is far the commonest order in German (c. 70% according to several authors, cf. a.o Lightfoot 1997 quoted in Frey 2006). As 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. However, according to Frey, not only the subject 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, Formal Movement is a property typical of V2 languages, and is satisfied by any constituent occupying the highest position below Fin. On the other hand, TAB is not restricted to V2 languages, and many other languages, if not all of them, have it. It is important to note that in V2 languages, either Formal Movement or TAB is able to satisfy the linear V2 requirement. According to Frey, there can thus be only one EPP feature in the C-domain.

Coming back now to the Romance languages like Classical Portuguese, we suggest that the similarity and the difference with Germanic languages may lie in the following point: *Classical Portuguese has V-to-C(Fin) but lacks Formal Movement.* Constituent fronting is therefore only forced by True A-Bar movement, under discursive pressure. This would straightforwardly account for the existence of V1. As for the frequency of V3, we will comment on it briefly in the Concluding Remarks.

In addition, and very interestingly, this analysis also explains a fact that up to now had remained unnoticed about the V2 orders themselves: the XV order in which X is not a subject is higher in Classical Portuguese than what is reported for German.

In order to understand this, it is important to observe that when we compare V2 orders in German and Classical Portuguese, we must keep in mind that one of them is a null subject language – so, even if the data presented above (cf. Fig. 2) does show that the frequency of VS is globally higher than the frequency of SV in the 16th and 17th centuries, it may not be sufficient to simply compare the rates of SV and VS, i.e. the possible positions for *lexical* subjects. In a language like German, any fronting of a non-subject constituent will result in VS (XVS); but in a null-subject language, when a non-subject constituent is moved to pre-verbal position, the subject, of course, will not necessarily remain post-verbal – 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 et al. (2010), as can be seen in Fig. 5a and 5b below. Figure 5a shows that the rates of V2 in Classical Portuguese are high even if we take into consideration V1 and V3 orders:

*Figure 5a: Evolution of V1, V2 and V3 in the diachrony of Portuguese* *(Cavalcante et al. 2010)*

A more interesting fact can be noticed if V2 sentences in which the preverbal phrase is not a subject are separated from V2 sentences that are superficially SV (as Figure 5b below does):
Figure 5b: Evolution of V1, V2 (SV and non-SV) and V3 in the diachrony of Portuguese (Cavalcante et al. 2010)

As Figure 5b shows, in the 16th and 17th centuries, the rates of XV are much higher than the rates of SV – respectively, 43% and 45%. In the 18th and 19th century, while the percentage of SV rises to 46% and 56% respectively, the XV orders (where X is not a subject) drops to 18% and 17%. To sum up: in V2 orders in general, Classical Portuguese displays more sentences in which the pre-verbal phrase is not a subject than sentences in which the pre-verbal phrase is a subject. It is worth emphasizing that this proportion (43-45% non-subject XV) is higher than what has been reported for German (30% non-subject XVs). If we look at individual authors, the XV data can also be very interesting, as 16th – 17th century authors with high rates of SV have even higher rates of XV: Pinto, for example, presents 51% of XV (against 20% SV, 28% V3 and 1% V1), according to Cavalcante et al (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 Classical Portuguese shows a definite preference for features that characterize V2 languages: the VS order (which is quite high – 21%-35%), and the XV order. In fact, we consider that the evidence for V2 in this data is very robust.

There is a final crucial point: if we are right in our hypothesis that Classical Portuguese is a grammar in which the pre-verbal phrases are fronted only for discursive/stylistic reasons, we expect a high variation among authors. This is exactly what we find, as we have shown in Section 2 (Table 1 and Figure 3), when we pointed out that the VS/SV variation among 16th – 17th century authors is considerably wider than the variation among Modern authors.

The natural question is then whether there is a relationship between the null subject property and the lack of formal movement. Antonelli (2011) argues that in CIP, 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 Classical Portuguese is a consequence of pro-drop. x

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

3.3. The change from Classical to Modern European Portuguese

As initially observed by Paixão de Sousa (2004), the surpassing of VS by SV is registered in the texts at the same time as the increase of the enclitic position of clitic pronouns. Fig. 6
below shows that the decrease of VS and the increase of SV-cl go hand by hand in the corpus of the present research as well:

**Figure 6:** Comparative evolution of the relative position of lexical subjects and enclisis, in clauses with pre-verbal subjects (VS/SV and S-Vcl/S-clV)

Up to 1700, SV-cl clauses are very marked constructions. We have seen above that, with the exception of Vieira’s Sermons, enclisis ranges from 0% to 18% in matrix SV sentences in the authors of the 16th and 17th century. In the authors born in the first half of the 18th century, we see that enclisis with pre-verbal subjects occurs at frequencies of 17% (Cavaleiro, b. 1702), 28% (Aires, b. 1705), and 44% in Marquesa de Alorna, born in 1750. After that, the frequency of SV-cl will reach almost 100% percent in the last author of our Corpus, Ramalho Ortigão, born in 1836. What we see in Aires and Cavaleiro seems therefore to be the beginning of a curve of change.

The hypothesis we put forth here to explain this is that the increase in enclisis with pre-verbal subjects (as well as with other pre-verbal phrases, although at a lower rate – cf. Galves et al. 2005), is the effect of a phonological change that affected the pronunciation of Portuguese at some point of the 17th century (cf. Frota et al. 2012 for an analysis based on the Tycho Brahe Corpus), and favored enclitic placement. This change, which was initially noted because it affected the pronunciation of pre-tonic vowels, is registered for the first time in a Grammaire Portugaise published in Paris in 1675 (cf. Teyssier, 1980). As argued by Revah (1952) “this is the more serious modification of the Portuguese pronunciation after the 16th century because it modified the very structure of words”. The 19th century Portuguese phonetician Gonçalves Vianna also 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: "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). Another property that can be inferred from the modern language (cf. Sandalo et al.
(2006 on secondary stress) to have emerged in this change, is the tendency to align stress and word boundary. This obviously creates a strong restriction on pre-verbal clitics.

The increase of enclitic placement has the effect of making the order Subject V-cl unmarked. Hinterholzl (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 a new prosodic pattern must arise with V-cl, one in which the pre-verbal subject is no longer associated with an independent intonational contour. This has two consequences: 1) the reinforcement of the interpretation of the pre-verbal subject as a non-contrastive topic, and 2) the change in the restriction on clitic-first. If the subject and the verb are in the same intonational phrase, the restriction on clitic-first can no longer be derived from the Tobler Mussafia law (cf. note iii). Leaving aside the nature of the new restriction (see Galves and Sandalo, 2012 for a proposal), in the authors of the 18th century, SVcl is no longer associated with contrastivity, as we now argue.

As already mentioned, Cavaleiro (1702) and Aires (1705) are likely to represent the beginning of a curve of change. However, this is not evident from the purely quantitative 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 17th century (52%). Again, a closer look at the data is necessary. The comparison with Vieira's Sermons is very instructive in this respect, because, as we saw above, the totality of cases of enclisis in the Sermons correspond to sentences in which the pre-verbal phrase is interpreted as contrastive topic. In contrast, when the pre-verbal phrase is a non-contrastive topic, proclisis is the exception-less rule.

In Cavaleiro, on the other hand, there is at least one case in which the pre-verbal subject of an enclitic sentence is clearly not contrastive.

(12) Estes meios empregavam se de duas sortes e tinham dois nomes.  
Those means used=SE of two kinds and had two names

‘Those means were used in two ways and had two names.’ (Cavaleiro, 1702)

In (12), “Estes meios”, these means, is a continuing topic, which shows that pre-verbal subjects with enclisis can be assigned a neutral interpretation. The analysis of the correspondence of Marquesa de Alorna (born in 1750) shows that this is a tendency for the generations to follow. In (13)-(14) we see that both proclisis and enclisis can be used with a subject whose interpretation is that of a continuing topic.

(13) Esta reflexão lhe basta para saber quem deve aceitar ou recusar quem must-3PS-PRES accept or refuse

‘This reflection suffices in order for you to know who to accept or to refuse’, (Alorna, 1750)

(14) Esta resposta aclarou-me... This answer clarify.3PS.PAST=me…

‘This answer enlightened me…’, (Alorna, 1750)
The variation in (13)-(14) nicely illustrates the competition between Classical Portuguese and Modern European Portuguese. (13) is representative of the former, with the pre-verbal phrase in a topic position prosodically integrated to the rest of the clause and the clitic, consequently, proclitic. (14), by contrast, is produced by the latter, with the pre-verbal phrase in subject position and obligatory enclisis. Neither of them could be associated with the alternative grammar, since proclisis is impossible in this context in Modern European Portuguese, and enclisis could not appear in this discursive context in Classical Portuguese.

We can also detect a change in the syntax of subjects in the first authors of the 18th century by observing the appearance of a novel construction in Cavaleiro (b. 1702) and Aires (b. 1705): the use of enclisis in V3 sentences in which the subject is in the middle position between some other phrase and the verb – XSV-cl:

(15) e verdadeiramente estes dois homens sós divertem me
and truly these two men alone amuse-3PP-PR=me

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)

(16) por isso os nossos afectos mudam-se
for this the our affects change-3PP-PR=themselves

‘for this reason our sentiments change’, (Aires, 1705)

As Galves and Paixão de Sousa (2005) had observed with a partially different corpus, in the 16th and 17th centuries, V3 sentences with enclisis are possible, but marginal – and they never present the order XSV, but only XXV and SXV. \(\text{xiii}\) Given the analysis of enclisis as a V1 phenomenon adopted in that paper, this fact indicates that when two external topics precede the verb, and the subject is one of them, the subject must occupy the more external position. The fact that the reverse order begins to appear in the authors born from 1700 on strongly suggests that a new position for the subject was made available, which means that the SVO grammar is already in use in texts.

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, grammar is in competition with the old, V2, grammar in the texts of the generations born after 1700. As regards clitic placement, the new patterns will evolve gradually – the modern, categorical use of enclisis in SV will be almost reached in the 19th century texts. As regards VS, the modern pattern is already instantiated in the 18th century. As we saw in section 1, there is a steep decline in the rates of VS in the first generation of authors born after 1700. The range of around 10% post-verbal subjects in those texts corresponds to postpositions that are produced by current Modern European Portuguese, and which, according to the literature (Ambar, 1992; Costa, 2004, a.o.), correspond either to unaccusatives or to the interpretation of the subject as focus and “residual” VS (such as inversions with affective operators)\(^\text{xiv}\).

4. CONCLUDING REMARKS

In this paper, we have proposed an alternative view of the loss of V2 (or, more accurately, the loss of V-to-C) in Classical Portuguese, which derives it from a change in the prosodic basis of the informational 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 co-existed with properties typical of V2 languages (VS and XV) over centuries.

More generally, we argue that what has been dubbed as ‘instability’ of V2 plus pro-drop in the literature derives from the fact that pro-drop V2 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 consequence on the perception learners have of the syntactic structure. In this sense, our analysis is close to the pioneer work on Old French by Marianne Adams who argued that the loss of V2 in this language was due to the loss of the Germanic stress.\textsuperscript{xv}

Here, we argued that the history of European Portuguese provides us with one more case study illustrating prosody-driven change.\textsuperscript{xvi}

It must be noted that the approach defended here allows us to reconcile somehow the two antagonistic lines of analysis about the V2 phenomenon in Old Romance languages.\textsuperscript{xvii} On 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 with 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 in this paper, we emphasize that V2 order in ClP depends exclusively on discursive requirements, in contrast with Germanic languages in which V2 is in most of the cases the result of a purely formal movement. This claim is totally compatible with Sitaridou’s affirmation above. However, from our point of view, this does not entail that the structure of main clauses in a language like Classical Portuguese 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 of VS and of XV empirically supports the conclusion that it does.\textsuperscript{xviii}

The consequence of this analysis amounts to shifting the difference between Germanic and V2 Romance languages from \textit{structural vs. linear V2} to \textit{formal-plus-discursive vs. discursive-only movement to Spec/CP}. This accounts for the frequency of V1 in the Romance languages, since a phrase only moves to the pre-verbal position when there is a discursive motivation for this. With respect to V3, an auxiliary hypothesis is needed. We tentatively propose that it has to do with different positions of the verb in the CP layer. Frascarelli & Hinterholzl (2007) argues 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 pre-verbal position in the latter, in contrast to the former. Additionally, we derive from this discursive-only movement to Spec/CP the historical fact that V2 in Romance is more prone to change than V2 in Germanic languages.

Finally, in Section 3, we have shown that in the texts written by authors born in the 18th century, there is evidence of the competition of two grammars: the old grammar of Classical Portuguese, and the new grammar of Modern European Portuguese. There are two alternative ways of understanding such a competition. 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 Section 3.1, the author argues that the acquisition process consists in the competition between grammars and that, in the presence of a heterogeneous environment, children can acquire two grammars, each one associated 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 2002:132-133). This model of language 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 children, when confronted to ambiguous data, to select the grammar that generate 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 (cf. for instance Kroch 2001), the variation in texts is not taken as being produced by the existence of two different grammars in the speaker’s mind as the result of his acquisition process, but rather as the socio-linguistic effect in written texts of former grammatical changes.  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 Classical and Modern European Portuguese relies on this view. We have argued that the steep change in the position of subjects was the effect of the reanalysis of SV sentences as having an underlying SVO order, in which the subject occupied no longer a topic position. We have also suggested that the change in clitic-placement plaid a crucial role in such a 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 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 of the change in subject position and the change in clitic placement in Portuguese: as can be seen taking together Figures 2, 3, 4 and 6, the rate of change in clitic-placement is different from the rate of change in the position of subjects: while the former happens in the time of a generation, the latter takes almost two centuries to reach completion. In fact, given the data presented in this paper, 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 paper to discuss in details the exact nature of the underlying changes involved. 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 an extremely salient phenomenon, to which speakers are sensitive, and to the additional fact that, once the position of pre-verbal subject has changed, the proclitic or enclitic position of the pronoun ceases to be associated with distinct prosodic patterns, and therefore loses any discursive effect on the interpretation of the sentence. In this situation, speakers cease to discursively discriminate proclisis and enclisis - but, as proclisis was the more frequent choice in the older generations, they continue to use it, and only progressively introduce in the written language the innovative form, which is enclisis. The view of grammar competition in texts as a result from previous grammar changes, therefore, is crucial to the conclusions we draw from the data on Classical Portuguese in this paper.

References


Frascarelli, Mara; Hinterhölzl, Roland. 2007. "Types of topics in German and Italian". In S. Winkler & K. Schwabe (Eds.), *On information structure, meaning and form* Amsterdam/Philadelphia: John Benjamins, pp. 87–116.

Frota, Sonia; Galves, Charlotte; Vigário, Marina; Gonzalez-Lopez, Verónica; Abaurre, Maria Bernadete. 2012 "The phonology of rhythm from Classical to Modern Portuguese". *Journal of Historical Linguistics* 2, pp. 173-207.


Révah, Israel. 1952. "L’évolution de la prononciation au Portugal et au Brésil du 16ème siècle à nos jours", in *Anais do primeiro congresso brasileiro da língua falada no teatro*.


Charlotte Galves
The Tycho Brahe Corpus is formed by texts written by authors born between 1380 and 1836, and currently comprises 57 texts (of which 16 are available with syntactic annotation, and 15 with corrected part-of-speech annotation). The Corpus is freely available for research and download at http://www.tycho.iel.unicamp.br/~tycho/corpus.


The data were extracted from non-dependent clauses, and concern all the verbs except 'ser' (to be). Because of their being categorically VS, we dropped from the quantification all the occurrences of parenthetical clauses following direct speech as in (i):

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

iii Galves, Britto & Paixão de Sousa’s analysis can easily be reformulated in terms of Rizzi’s (1997) expanded CP. In fact, the external/internal divide should be understood more in prosodic terms than in syntactic terms. According to Frascarelli and Hinterhölzl (2007), contrastive topics are high in the hierarchy of syntactic CP nodes and receive an independent intonational contour. Familiar topics are low and are part of the intonational contour of the rest of the sentence. This is coherent with the claim that what governs clitic-placement in Classical Portuguese is the Tobler-Mussafia law, which prevents a clitic from standing at the beginning of an intonational phrase.

iv Note also, that the same reasoning would lead to expect that in a heterogenous environment composed of V2 and SVO speakers, the V2 language would take over the SVO language, since acquirers would end up with more non-ambiguous evidence for V2 than for SVO. May be this was the case when Germanic tribes invaded the Romance Empire in a period in which Proto-Romance languages were already SVO.

v In Yang’s framework, one could possibly argue that Romance V2 languages are a stable combination of V2 with SVO grammars. Such a claim would, again, leave unexplained why, at some point, instability arises. Instead, we argue in 3.2.2 that V3 is compatible with the V2 grammar instantiated by Classical Portuguese, and possibly other Romance languages.
The presentation here is based on Frey (2006).

Additionally, Frey proposes that some adverbs can be inserted directly in Spec/CP. This is not relevant for our analysis.

This was first argued by Antonelli (2011). For further arguments see Galves and Gibrail (2013).

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

Alternatively, one could argue that it is satisfied by the phi-features of the verb. Thanks to Tony Kroch to have drawn our attention onto this point.


Frota et al. (2012) look at the distribution of the words in 16 texts of 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 17th century on. They interpret this fact as the integration in the language of new rhythmic properties, characteristic of stress-timed languages.

It must be noted that Galves and Paixão de Sousa (2005) find one case of XSV-cl, in Bernardes, born in 1644:

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

After the paper was published, the authors discovered that this sentence had not been appropriately analysed because of a modification of the punctuation performed by 20th century editor of the version used in the Tycho Brahe Corpus. The punctuation of the original edition is as follows:

Vendo tão rara, e verdadeira amizade el-rei Dionísio o mais velho, disse-lhes:
This strongly suggests that there is another analysis for this sentence, in which the DP “el-rei Dionísio” is not the pre-verbal subject of the tensed verb, but the post-verbal subject of the gerundive verb. If this analysis is right, there is no occurrence of the order XSV-cl in the Corpus.

Eide (2010:150) shows that there is a difference in the decrease rate of VS according to whether S is new or old information: VS where S is new information remains at higher rates up to the end of the corpus in this study, which extends up to the end of the 20th century.

Prosody also plays a crucial role in the analysis of the loss of V2 in French in Kroch (2001). 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.”

Following Galves and Galves (1995) and Galves et al. (2005), Eide (2010) also assigns prosody a crucial role in the loss of V2 in Portuguese.

cf. also Galves and Gibrail (2013) for an argumentation based on the position and interpretation of subjects in transitive sentences.

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 [our emphasis, CG; MCPS].” 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 sociolinguistic methods to observe unreflecting speech. Such studies do not at present exist, in part because syntactic 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 predictions is a major task for the future of diachronic syntax.”

We assume that the reanalysis of the position of pre-verbal subjects involves a shift in the features associated to C and T. As for clitic placement, we adopt Galves & Sandalo, (2012)’s claim that enclisis is derived by a post-syntactic 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 Modern European Portuguese.