1. Introduction

In this paper, we describe the loss of V2 in the history of a pro-drop Romance language, European Portuguese. We bring data extracted from a parsed Corpus now available to study the relevant phase of the history of this language, the Tycho Brahe Corpus, which consists of texts written in Portuguese by authors born between 1380 and 1850. This corresponds to what has been called, in the tradition of Historical Portuguese, Middle Portuguese, Classical Portuguese, and the beginning of Modern Portuguese. Our focus is on the change from Classical Portuguese to Modern European Portuguese - following the tradition, we call “Classical Portuguese” the language instantiated in texts written by authors born between the 16th and 18th centuries. Previous studies have shown that this language corresponds to a grammar which is different both from Old Portuguese and Modern European Portuguese (cf. Galves et al. 2006). Classical Portuguese is a pro-drop, V2 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).

On one side, therefore, we are offering the empirical evidence pointed out by Yang (2000) as necessary for the full understanding of the problem of the instability of V2 in Romance languages. We shall adopt the framework of Yang (2002) to examine the grammatical change from V2 to SVO in the history of Portuguese in the light of the following observation: the generation of Portuguese authors in our
corpus born at the beginning of the 18th century is representative of what Yang calls the "puncturing of homogeneity". In the texts written by this generation, V3 sentences (XP-Subject-V) become as frequent as VS (XP-V-Subject) sentences. According to Yang, this creates conditions for the development of ambiguity between an SVO (V3) and a V2 (VS) grammar.

On the other side, as we will show, the data on Classical Portuguese indicates the need for a review of Yang's proposal. According to Yang (2002), the emergence of heterogeneity in V2 Romance languages springs from the combination of V2 and null subjects, which would be "intrinsically unstable". However, we shall argue that this is not a sufficient account for the loss of V2 in Portuguese, since this 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. 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 (see the development of XSV-cl). 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.

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), and in section 3 further on we analyze this data in the framework of Yang (2002)'s model. Finally, in the concluding remarks, we suggest that, as far as change is concerned, our data bring empirical evidence for an alternative view of grammar competition, closer to what has been proposed by Kroch (1994, 2001). In this view, the grammar
competition we observe in texts is not the cause but rather the result of the change, and has to do with the conservative nature of written texts.

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 an annotated corpus - comprising XXXX declarative main clauses in texts written by 15 authors born between 1510 and 1802.ii

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. However, like with other Old Romance Languages, "V2" in Classical Portuguese must be understood in a particular context: that of a null subject system, i.e. a system in which superficial verb-second coexists liberally with verb-first and verb-third orders. Given our central aim - to model the change between such a system and a canonical SV system, that of Modern European Portuguese - we will, in the present paper focus on the possible expressions of subjects in Classical Portuguese: lexical (VS or SV) and null. As for the varied possible positions for the verb in Classical Portuguese, we shall make reference, where relevant, to Cavalcante et al. (2010) and Antonelli (2011). Given, also, the nature of our main hypothesis - according to which clitic placement plays a key role in the change - we
shall also 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 to Modern European Portuguese, as Fig. 2.1 illustrates.

**Fig.2.1:** Relative Position of Lexical Subjects in matrix clauses(VS/ SV)

We take the high frequency of VS orders as one of the most important pieces of evidence for a V2 system in Classical Portuguese. The data on subject-postposition includes both verb-initial constructions and constructions on which a pre-verbal element other
than the subject appears in the pre-verbal position - X V (S), be it X V S (X fronted, post-verbal Subject), or X V (X fronted, null subject), both quite frequent in these texts. We take both X V S and X V as instances of general "X V", i.e., of "non-S V" 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 discoursive-prominent elements. We are not concerned, in this paper, with the exact nature of this position (see Antonelli, 2011, for a concrete proposal in the framework of Rizzi’s cartography).

The relevant point is that Classical Portuguese is not an S V 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, S V (1), V S (2), and null subjects (3):

(1) Christo Senhor nosso, disse a seus Discípulos, que o segredo d’aquelle dia
Christ Lord our, said.3PS.PAST to his Disciples, that the secret of that day
é reservado só ao Padre. 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) Mártires os chamavam os companheiros.
Martirs them= called.3PP.PAST the fellows
‘Their fellows called them martyrs’ (Sousa, 1556)

(3) e o que não podiam trazer ou o despedaçavam,
and the what not could.3PP.PAST bring or it= smash.3PP.PAST, ou lhe punham o fogo.
or to-it= put.3PP.PAST the fire.
‘and that which they could not carry, either they smashed it, or they put fire to it’. (G alhegos, 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). Fig. 2.2 below shows the distribution of the possible patterns for subjects in main clauses between 1500 and 1850:

**Fig. 2.2: 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 the groups of texts representative of the 16th and 17th centuries, the proportion of VS over SV and Null subjects varies, but 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 47% against 15% in the 18th century, and 42% against 12% 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 contemporaneous texts in the corpus - and that this variation also contrasts the two periods under study (pre-18th century and post-18th century). Table 2.1 below shows this, depicting the data for SV and VS in each of the eleven texts of our corpus.

Table 2.1: Evolution of the expression of subjects, 16th to 19th centuries

<table>
<thead>
<tr>
<th>Year</th>
<th>Pinto</th>
<th>Sousa</th>
<th>Galheg.</th>
<th>Vieira</th>
<th>Céu</th>
<th>Barros</th>
<th>Cavaleiro</th>
<th>Aires</th>
<th>Marq.</th>
<th>Alorna</th>
<th>Ortigão</th>
</tr>
</thead>
<tbody>
<tr>
<td>1510</td>
<td>0.24</td>
<td>0.13</td>
<td>0.21</td>
<td>0.34</td>
<td>0.17</td>
<td>0.15</td>
<td>0.35</td>
<td>0.57</td>
<td>0.38</td>
<td>0.45</td>
<td>0.37</td>
</tr>
<tr>
<td>1556</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1597</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1604</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1658</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1702</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1705</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1750</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1802</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1836</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></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 are more interested here in the variations concerning the position of lexical subjects. In this regard, notice that it is in the authors born after 1700 that we start to see a consistent pattern: in all the texts written by those authors, the frequency of SV, however variable, is always higher than the frequency of VS. Compare this to the previous period (1500-1600), when the proportions of each pattern vary considerably for each text, but VS surpasses SV in most cases.

There are two texts in which the proportions seem not to fit in with the general pattern for their period: Vieira's Sermons, with a proportion of 34% SV, 33% VS and 33% null subjects; and Pinto's Peregrinação, with 24% SV, 14% VS and 62% null subjects. We interpret this in the light of the style in which the texts were written - and we are not surprised to see oscillations in this regard, considering that word order in V2 languages is closely related to discursive constraints. 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 to Modern European Portuguese is that enclisis with subjects increases sharply after the 18th century. Fig. 2.3 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):
Fig. 2.3: 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 before the 18th century reveal rates of 0% to 9% of enclisis in SV sentences, while texts written by authors born after that exhibit rates of 19% to 94% enclisis in this context. This would be so, if it were not for one 17th century text: Vieira's Sermons, with 55% 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. Further on, we turn to general pre-verbal subjects, showing some data from our other "exceptional" text (Peregrinação, by Pinto, with an unusually high rate of SV, as we saw).

Galves (2002) and Galves et al. (2005) have argued that SV with enclisis in Classical Portuguese texts correspond to contrastive
topicalizations, which can be seen in the examples below, from Vieira's Sermons.

(4) Elles *conheciam*-se, como homens, Christo *conhecia*-os, como Deus.

They *know.*3PP.PAST=themselves, as men, 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; os homens *julgam*-nos a nós por si. (Vieira, 1608)

God *judge.*3PS.PRES=us to us for us; men *judge.*3PP.PRES=us to us for themselves.
'God judges us for ourselves; men judge us for themselves.'

As it was shown in that 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, Galves et al. (2005) pointed out that in the Sermons, proclisis is used in distinct discourse environments to enclisis - not only in focalizations, as claimed by Martins (1994), but also in sentences involving non-contrastive topics, as in the examples below.

(6) O Evangelho o *diz*: Erunt signa in sole, et luna, et stellis.

'The Gospels it= say.*3PS.PRES:....
'The Gospels says it: ...'

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

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

The different topics in the examples above 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 topics" correspond 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.

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.
2.3). Let us examine the following examples of SV with proclisis in this text:

(8) O capitão-mor lhe respondeu que
The captain-major to.him= answer.3PS.PAST that
os embaixadores tinham seguro para suas pessoas. (Pinto, 1510)
the ambassadors have.3PP.PAST insurance for their persons
'The captain-major answered him that the ambassadors had personal insurance.'

(9) O brâmene lhe deu por isso seus agradecimentos, (Pinto, 1510)
The braman to.him= give.3PS.PAST for that his thanks,
'The braman thanked him for that,'

As becomes clear when one examines the context from which the sentences above have been extracted, they are part of a narrative sequence which 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 a similar pattern; 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 commit.3PS.PAST then to.burn=him the gaol,
[Pinto, 1510]
'our captain-major proceeded then to burn his gaol'

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 $X V (S)$ sequences - i.e., $X V$, with null subjects; or
$X V S$, 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... [Pinto, 1510]
   'To the merchant who brought me, Pero de Faria ordered that sixty
cruzades be given...'

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 $V S$ - more specifically, $X V S$ - can be accounted for in
Classical Portuguese: instances of V2-fronting of non-subjects. The
proportion of $S V$ and $V S$, 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 2.1, 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 two possible structural positions for pre-verbal constituents - one of them "external" to the clause, the other "internal". Both positions host topicalized elements, with distinct information-structure correlates: the external position hosts contrastive topics, the internal position hosts other topics and prominent elements - and in both cases, this includes SV:

Classical Portuguese:
SV = Topic Verb i.e.,
(a) [Contrastive Topic] # Verb
(b) # [Topic-Focus] Verb

Sentences with clitics are key data for this structure, as the restriction on clitic-first correlates S-V-clitic with a marked prosodic pattern, corresponding to structure (a) above - we will explore this further in section 3. 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 in the framework of Yang (2002).

3. The change from Classical to Modern European Portuguese

This section is organized as follows. In 3.1, we present Yang's model. In 3.2, we show how the general model of grammar competition can shed light on the change from Classical Portuguese to Modern European Portuguese. As for Yang's interpretation of the loss of V2 in Old Romance Languages, however, we argue that it does not receive empirical support from the history of European Portuguese.

3.1 Yang's Model

Yang's (2002) model of language change derives from a model of language acquisition viewed as grammar competition. In this model, 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 $G_i$ is paired with a weight $p$ which can be viewed as the measure of prominence of $G_i$ 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).

This 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 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 (X VSO, O VS), and evidence for SVO is V3 (SXVO, X SVO). 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%. In languages like English, however, only 10% of all sentences are V3 patterns. 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 accounts 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 diminished in the total dataset because of the existence of null-subject sentences. 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, we find a frequency of 5-7-18% according to 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 OF V2 grammar. Yang concludes that the combination of pro-drop and V2 is "intrinsically instable and will necessarily give away to an SVO (plus pro-drop) grammar" (Yang, 2002: 137).

In the next section, we examine 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.

3.2 The change from ClP to EP

3.2.1 Grammar competition in the 1700

As we have shown in Section 2, our data evidences a sharp increase of pre-verbal subjects between the last author of the 17th century (André de Barros, born in 1675) and the first author of the 18th century (Cavaleiro de Oliveira, born in 1702). From the point of view of Yang's analysis, the advantage of the V2 grammar over the SVO one is computed from the difference between the frequency of VS and the frequency of V3. According to Cavalcante et al. (2010), who work with the same corpus as in the present study, V3 sentences occur at the rate of 13%, and 10% in the 16th and 17th respectively. This is much lower than the frequency of VS, even if we take null subjects into consideration. It therefore seems that up to 1700, the V2 grammar was very strongly evidenced in the data, and there was no reason why children should have also selected the SVO grammar. We return later to the analysis of V3 in a V2 grammar.

The question is therefore what caused this drastic change in the space of a generation. Note that since Classical Portuguese is a
grammar in which the pre-verbal phrases are fronted for discursive/ stylistic reasons, we expect to find variation among texts in the kind of pre-verbal constituents. The strong increase in pre-verbal subjects could therefore be understood, at its beginning, as the effect of the topicalization of subjects to the detriment of the fronting of other constituents (in the case of Aires (1705), we see that this happens also to the detriment of null subjects). 

One possible hypothesis to interpret this set of data would be to claim that, at its beginning, the increase of pre-verbal subjects and the concomitant decrease in VS is the effect of a higher frequency of subject topicalization, inside the V2 grammar. Cavaleiro de Oliveira and Aires could be viewed as reflecting a general tendency of their time. In the framework of Yang's analysis, the language produced by their generation would have the required properties to provoke a grammatical change in the next generations: the rate of post-verbal subjects, becomes lower than the rate of V3, giving a positive weight to the SVO grammar. This means that children would be led to select both grammars.

Putting things this way leads us to ask whether in Cavaleiro and Aires, there is a quantitative change only or if there is evidence that they themselves already instantiate grammar competition between V2 and SVO. Indeed, in their texts, we find empirical evidence of both the old and the new grammar. The facts supporting this claim have to do with clitic-placement.

First, Cavaleiro and Aires both present a novelty in the history of the language, which is the use of enclisis in V3 sentences in which the subject is in the middle position between some other phrase and the verb:

(12) e verdadeiramente estes dois homens sós divertem me mais que a comédia toda junta.
And truly these two men alone amuse me more than the comedy all together (Cavaleiro, 1702)
‘and truly these two men alone amuse me more than all the comedy’

(13) por isso os nossos afectos mudam-se, (Aires, 1705)
for this the our affects change-3PP-PR=themselves
‘for this reason our sentiments change’

As Galves and Paixão de Sousa (2005) had already observed with a partially different corpus, in the 16th and 17th centuries, although enclisis appears in V3 sentences (at a rate similar to what is observed in V2 sentences), it never shows up with the order XSV, but only with XXV and SXV. 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.

Second, we find in Cavaleiro and Aires a much higher rate of enclisis in V2 sentences with pre-verbal subject than in the previous generations: 19% and 30% respectively. These are very high rates when compared with what we find in the 16th and 17th century texts, with the exception of Vieira's sermons, as already discussed. The comparison with Vieira's sermons is very instructive because, as we mentioned 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 exceptionless rule.

In Cavaleiro, there is at least one case in which the pre-verbal subject of an enclitic sentence is clearly not contrastive.
In (14), “Estes meios”, these means, is a continuing topic, which shows that pre-verbal subjects with enclisis can be assigned a neutral interpretation. We return to this fact. For the time being, we take the two novelties described so far as evidence that a new, SVO, grammar is in competition with the old, V2, grammar in the texts of the generation born at the beginning of the 18th century.

In the next section, we return to the question of the conditions under which V2 was lost in Classical Portuguese. We argue, on the basis of the data provided by our Corpus, that a model that only takes the frequency of patterns of order into consideration, as in Yang (2002), cannot account for this particular change.

3.2.2 Problems for/ with Yang’s analysis

In this section, we point out problems in Yang’s reasoning, as far as Classical Portuguese is concerned. As mentioned by many authors, in Old Romance Languages, V3 is much more frequent than in Germanic V2 languages (cf. Cruschina and Sitaridou (2010) a.o.). For Classical Portuguese, Galves and Paixão de Sousa (2005) argue that the V3 order derives from the existence of (at least) two topic positions, one internal (the V2 one) and the other external. V3 order therefore harmoniously co-exists with V2, and V1, and even with some V4 (cf. Gibrail 2010). This means that, when confronted with a V3 sentence, speakers of ClP know that the phrase closest to the verb is in the internal TOP position and the more peripheral one is in the external TOP position. Galves et al. (2005) also argue that certain V2 orders correspond to a structure in which the pre-verbal phrase is not
in the internal TOP, but in the external one (see also Salvi 1991 for a similar claim). This is the case for V2 sentences in which the clitic is post-verbal, as in (4-5) above. In those sentences, as argued in the last section, the pre-verbal phrase is interpreted as a contrastive topic. In this case, crucially, the pre-verbal phrase is assigned an autonomous intonational contour. This contour plays an important role in the characterization of the pre-verbal phrase as a topic.

From this point of view, V3 sentences are not incompatible with the type of V2 grammar instantiated by Old Romance languages, including Classical Portuguese. We have seen that their frequency is around 10% in the 17th and 18th century. According to Yang (2002), this is comparable to what is observed in an SVO language. At the same time, despite the existence of null subjects, we find a rate of VS typical of V2 languages. This kind of system strongly questions Yang’s model, which predicts that in this case, children will postulate two grammars. The problem for this analysis is the stability of the system over many centuries and its abrupt change at the turn of the 18th century.

The data from the Tycho Brahe Corpus, therefore, does not confirm the role of null subjects in the loss of V2. When we look at the distribution, by authors, of SV, VS and null subjects over time in Portuguese (cf. Fig. 2.2), we see that the frequency of null subjects does not substantially change. What we find in all the periods is significant variation across authors. The presence of null subjects does not prevent Classical Portuguese from robustly exhibiting its V2 properties, since, as repeatedly mentioned, VS occurs at a frequency comparable with what is found in Germanic languages.

Summarizing, Classical Portuguese brings evidence that there is no intrinsic instability in the conjunction of V2 properties and null subjects. Besides, we see that V3 and V2 can co-exist at rates associated in other languages with distinct grammars. Moreover, the rate of pre-verbal subject placement is variable, and in some texts, can be higher (cf. Pinto) or equivalent (cf. Vieira’s Sermons) to the rate of
post-verbal subjects, with no effect on the next generations. This is expected in a system in which pre-verbal subjects are topics, since topicalization depends on informational and stylistic factors. We also observed that the novelties found in the texts written by authors born at the very beginning of the 18th century concern not only the placement of subjects, but also the placement of clitics. This suggests that another factor is in play in the abrupt change observed. In the next section, we propose an alternative analysis in which prosodic factors play a decisive role in the grammatical change.

3.2.3 The role of prosody in the change from Classical to Modern European Portuguese

As initially observed by Paixão de Sousa (2004), at the same time as SV surpasses VS, we also observe an increase of the enclitic position of clitic pronouns. In Cavaleiro and Aires, we see that enclisis with a pre-verbal subject occurs at a frequency of 19% and 30% respectively. Fig. 3.1 shows that the decreasing of VS and the increasing of SVcl go hand by hand.

Fig. 3.1: Comparative evolution of the relative position of lexical subjects and enclisis (VS/ SV and S-V cl/ S-clV)
As we have already seen above, up to 1700, SV-cl clauses occur at a very low frequency except in Vieira's Sermons. This shows that they are very marked constructions. The question that is raised by the increase of this kind of sentence at the beginning of the 18th century, is what initially provoked it. If it were the case that the higher rate found in Cavaleiro and Aires was the effect of stylistic requirements, it would have remained as isolated in its time as it was in Vieira's Sermons. But we see that after them, the frequency of SV-cl continues to raise, reaching 42% in Marquesa de Alorna, born in 1750, and almost 100% percent in the last author of our Corpus, Ramalho Ortigão, born in 1836. What we see in Aires and Cavaleiro is therefore the beginning of a curve of change. The hypothesis we put forth here is that the increase in enclisis with pre-verbal subjects as well as with other pre-verbal phrases, although at a lower rate, is the effect of a phonological change that affected the pronunciation of Portuguese at some point of the 17th century (cf. Frota et al. to appear 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. to appear). 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, p. XX ) 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-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), we can see that in the authors of the 18th century, SVcI is no longer associated with contrastivity. This was illustrated above with a sentence from Cavaleiro (born in 1702). We can observe the same fact in sentences drawn from the correspondence of Marquesa de Alorna (born in 1750). In (15)-(16) we see that both proclisis and enclisis can be used with a subject whose interpretation is that of a continuing topic.

(15) Esta reflexão lhe **basta** para ... .  
This reflection to.(you= suffice.3PS.PRES for ... .  
saber quem deve aceitar ou recusar .(Alorna, 1750)  
to.know who must-3PS-PRES accept or refuse  
‘This reflection suffices in order for you to know... who to accept or to refuse’.

(16) Esta resposta **aclarou**-me... (Alorna, 1750)  
This answer clarify.3PS.PAST =me ...  
‘This answer enlightened me...’

The variation in (15)-(16) nicely illustrates the competition between Classical Portuguese and Modern European Portuguese. (15) is produced by the former, with the pre-verbal phrase in internal/low topic position and the clitic, consequently, proclitic. (16), by contrast, is produced by the latter, with the pre-verbal phrase in subject position and obligatory enclisis. Both of them could not 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.

4. Concluding Remarks: choosing between two models of
In this paper, we have so far assumed Yang (2002)'s theory of acquisition and change, which allows for children to select two grammars as the result of their acquisition process. 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. Although Yang claims that the combination of two grammar remains stable in the speaker’s mind, he assumes that it is diachronically unstable. In the case of Portuguese, as shown by Figure 2.3, from the beginning of the 18th century, the syntax of clitic-placement gradually changes from the CIP pattern to the EP pattern, reaching completion more than one century later.

However, Yang's model can be questioned from the point of view of his theory of acquisition. It is far from being commonly accepted that children, given certain conditions, select two different grammars at the end of their acquisition process. In Yang (2002), empirical evidence for this claim comes mainly from the variation found in historical texts. But it must be noted that such variation can be accounted for as well by a weaker model of grammar competition, by which the competition would not be produced by the existence of two different grammars in the speaker’s mind, but would be the effect of competition in the texts between the unique grammar selected by the writers in their acquisition process, and fragments of old grammars, due to the conservative character of written texts (cf. Kroch, 2001). Second, 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 VS over centuries. Moreover, the change in the position of subjects occurs in the space of a generation, which is rather unexpected under Yang's
model, since we would expect to see the effect of the competition between the two grammars during a larger span of time. Finally, we proposed an alternative account for the loss of V2, viewed as the effect of a prosodic change affecting grammar selection. This is not a new idea, in particular with respect to the loss of V2. In the framework of generative grammar it was first argued by Adams (1987) that V2 disappeared in French due to the loss of the Germanic clause-initial stress. A similar idea is reformulated in Kroch and Santorini (2009). In this paper, we argue that the history of European Portuguese provides us with one more case study illustrating prosody-driven change.

5. 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; González-Lopez, Veronica; Abaurre, Maria Bernadete. to appear "The phonology of rhythm from Classical to Modern Portuguese". Journal of Historical Linguistics.


The parsed Corpus is composed of the following authors and works: Fernão Mendes Pinto (b. 1510) Peregrinação; Frei Luís de Sousa (b. 1556) A vida de Frei Bertolameu dos Mártires; Manuel Galhegos (b. 1598) Gazeta; Pe António 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; Marquês de Fronteira e Alorna (b. 1802) Memórias; Ramalho Ortigão (b. 1836) Cartas a Emília.

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.

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 tendency to topicalize subjects is not only expressed in V2, but also in V3 constructions. Antonelli (2011) observes that the increase of V3 at the beginning of the 18th century is accompanied by a modification of the distribution of the pre-verbal elements. In the 16th and 17th century, there is a balance between XSV, SXV and XXV. In half of the occurrences of V3 sentences, one of the pre-verbal phrases is a subject. From 1700 on, the frequency of pre-verbal subjects jumps to near 75%.

As argued by Ribeiro (1995), Old Portuguese (12th-15th centuries) was also a V2 plus Null Subject language.

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 oxytone 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.

The fact that the competition between enclisis and proclisis takes a much longer time to reach completion could be an argument in favor of a [model of competition in texts. Whereas people consider
SV a natural order, given the strong variation in subject position in the old grammar, they keep the memory that enclisis in V2 was a marked construction, and tend to resist the pressure of the vernacular when they write.