

A Comparative Diachronic Analysis of *Wh*-Questions in Brazilian and European Portuguese

Mayara Nicolau de Paula^a

ABSTRACT

Based on the Principles and Parameters framework (CHOMSKY, 1995), this article presents a diachronic analysis of European Portuguese (EP) *Wh*-questions and a comparison with the same structures in Brazilian Portuguese (BP). I also present a brief analysis of *Wh*-question patterns in sociolinguistics interviews recorded in two periods (years 1970/80 and 2010). The initial hypothesis, inferred from recent theoretical descriptions, was that WhVS order would be more frequent in EP and that WhSV, whenever attested, would be constrained by the presence of the cleft-structure. Taking into account the fact that EP is considered a consistent Null Subject Language (NSL), 1st and 2nd person subjects as well as anaphoric subjects would be preferably null. Therefore, EP would exhibit a different behavior from BP, which has become a WhSV system; the rare cases of SV are attested with unaccusative verbs and a lexical DP subject; and this change, as shown by the analyses reported here, runs parallel to the re-setting of the value of the Null Subject Parameter in BP. The sample analyzed for the diachronic study, comprising theater plays written across the 19th and the 20th centuries, is comparable to the Brazilian sample. For the interviews, I used the samples *NURC* and *Concordância* for PB and *Cordial-Sin* and *Concordância* for EP. The methodology to codify and submit the data to statistical treatment follows the variationist approach (TAGLIAMONTE, 2006; GUY AND ZILLES, 2007). The results show that EP prefers the WhV pattern, confirming its status of a NSL; as for overt subjects, WhVS is the preferred pattern; however, we can observe a slow decrease of VS in the last quarter of the 20th century, suggesting the implementation of a competition with WhSV, triggered by the introduction of clefting in the second half of the 19th century. Once introduced in the system, the cleft structure expands to all Wh patterns, which is confirmed in contemporary speech data.

KEYWORDS: *wh*-questions; VS/SV order; cleft structure; null subject parameter; European Portuguese-Brazilian Portuguese

RESUMO

Com base na Teoria de Princípios e Parâmetros (CHOMSKY, 1995), apresento uma análise diacrônica das interrogativas-Q do Português Europeu (PE) seguida de uma comparação com as mesmas estruturas no Português Brasileiro (PB). Paralelamente, faço uma breve análise dos padrões de interrogativas-Q em entrevistas sociolinguísticas gravadas em dois momentos (anos 1970/80 e 2010). A hipótese inicial, a partir de descrições recentes de base teórica, era a de que a ordem QVS no PE seria a mais frequente, enquanto a ordem QSV estaria sempre condicionada à presença da clivagem. Sujeitos de 1^a. e 2^a. pessoas bem como os anafóricos seriam preferencialmente nulos no PE, um sistema descrito como de sujeito nulo

^a Universidade Federal de Minas Gerais. E-mail: maynicolau@gmail.com

consistente. Nesse aspecto, o PE teria um comportamento diferente do PB, que perdeu a ordem QVS, hoje atestada apenas em estruturas com verbos inacusativos, desde que o sujeito seja um DP lexical, uma mudança paralela à remarcação do valor do Parâmetro do Sujeito Nulo no PB. A amostra analisada para o estudo diacrônico é constituída de peças portuguesas escritas ao longo dos séculos 19 e 20, comparável à amostra brasileira que nos serve de ponto de comparação. No caso das entrevistas sociolinguísticas, foram utilizadas as amostras *NURC* e *Concordância* para o PB, e *Cordial-Sin* e *Concordância* para o PE. A metodologia para o tratamento dos dados segue o modelo variacionista (TAGLIAMONTE, 2006; GUY E ZILLES, 2007). Os resultados mostram que o PE prefere o padrão QV, com sujeitos nulos; quanto aos sujeitos expressos, a ordem QVS é o padrão preferido; observamos, no entanto, uma curva descendente no último quartel do século 20, sugerindo o início de uma competição com QSV, desencadeada pela entrada da clivagem, a partir da segunda metade do século 19. Uma vez introduzida no sistema, a clivagem se expande para os três padrões de interrogativas-Q, o que é confirmado pelos dados da fala contemporânea.

PALAVRAS CHAVE: interrogativas Q; ordem QVS/ASV; clivagem; parâmetro do sujeito nulo;
Português Europeu and Português Brasileiro

Introduction

This paper describes the Subject (S) Verb (V) order in European Portuguese (EP) *Wh*-questions in matrix clauses, based on a sample of popular plays comprising the 19th and 20th centuries, and compares the results attested for EP (NICOLAU DE PAULA, 2016) with Duarte's (1992) findings for Brazilian Portuguese (BP), later refined by Pinheiro and Marins (2012), who included *Wh*-questions with null subjects using the same sample of plays used by Duarte. This would allow an examination of three possible patterns: WhVS, WhSV and WhV

The general hypothesis guiding this study is based on Duarte and Kato (2002), who suggested that the change in BP attested by Duarte (1992) – from WhVS to WhSV – was closely related to the rise in the use of overt referential pronominal subjects in all kinds of sentences (declaratives, yes/no questions and *Wh*-questions) in the same sample (cf. DUARTE, 1993). A comparison of both analyses allowed the hypothesis that rates of overt pronominal subjects in declaratives, which preceded the change towards SV in *Wh*-questions, would have functioned as a trigger for this change. The high rates of overt pronouns would lead a child to interpret a null subject in a WhV sentence as preposed to the verb Wh (S) V, while older generations's null subjects would be in a structure like WhV (S). Therefore, overt subjects and WhSV order would be closely related. Since EP is a consistent null subject system (see BARBOSA, 1995; DUARTE,

1995; HOLMBERG, 2010; LOBO, 2013, among others), one would expect a different behavior, with subjects preferably in WhV (S) pattern. WhSV pattern, whenever attested, would be constrained by the presence of the cleft structure, according to descriptions of standard EP (AMBAR, 1992; BRITO, DUARTE AND MATOS, 2003).

The paper is organized as follows: the next section presents Duarte's (1992) diachronic analysis of the change from WhVS to WhSV in BP, followed by Pinheiro and Marins's (2012) work, which confirms the co-relation suggested by Kato and Duarte (2002), investigating *Wh*-questions with null subjects in the same sample of Brazilian plays analyzed by Duarte (1992). I then review descriptions related to VS order in EP *Wh*-questions and present the diachronic sample collected for the present analysis. In the next section, both the diachronic and synchronic results for EP are presented and compared with BP. Finally, I discuss the differences between BP and EP *Wh*-questions with respect to the loss of null subjects in the former and the predominance of null subjects in the latter, the importance of the introduction of the cleft structure to trigger WhSV order in both varieties and its propagation to WhV and even WhVS questions in EP.

1. *Wh*-questions and VS–SV order in Brazilian Portuguese

Duarte's (1992) analysis of *Wh*-questions in BP popular plays suggested that the introduction of the cleft structure might have been the trigger of the loss of VS. Fig. 1 shows the rise of SV order across seven synchronies (the year of the play analyzed represents each synchrony).

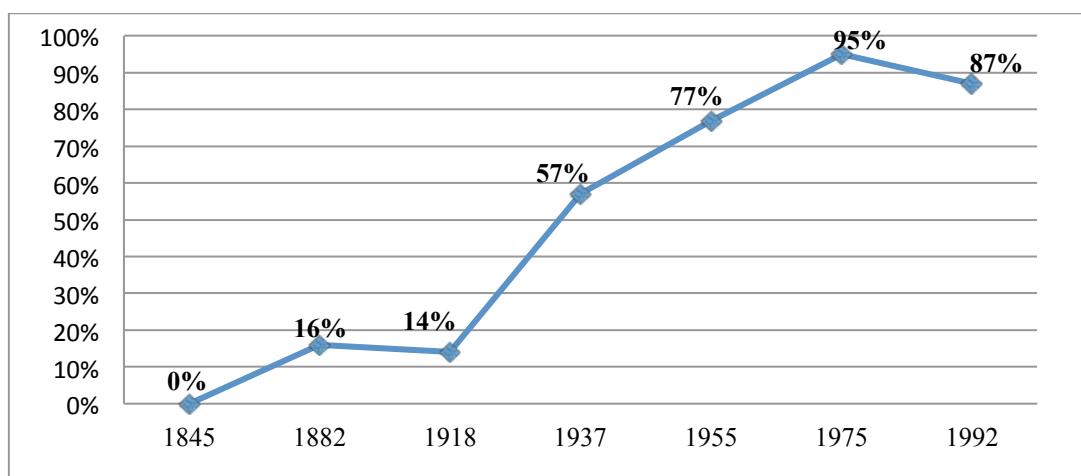


Fig. 1. WhSV (vs WhVS) across two centuries (adapt. from Duarte, 1992:41)

In the first half of the 19th Century (1845), the pattern attested is VS, as in (1):

- (1) O que **pensa tua filha** do nosso projeto? (1845 - BP)
 what think.3SG your daughter of-the our plan
 ‘What does your daughter think of our plans?’

Only in the 2nd half of the century are clefts introduced in the system, occurring in 4 instances showing WhSV order (2a) and one instance of WhVS (2b); in the same synchrony, two WhSV questions with no cleft (both with the interrogative “why”), instantiated in (2c), are attested:

- (2) a. O que **é que tu tens** nesta barriga? (1882 - BP)
 what is that you have.2PS in-this belly
 ‘What do you have in this belly?’
- b. E onde **é que foi a mulher do Seabra?** (1882 - BP)
 and where is that went the wife of-the Seabra
 ‘And where did Seabra’s wife go?’
- c. E para que **você quer ser** ministro, seu Chico? (1882 - BP)
 and for what you want to-be minister, Mr. Chico
 ‘And why do you want to be a Minister, Mr. Chico?’

The low frequency of WhSV continues in the play representing 1918, with 6 instances (out of 29 questions), 3 with cleft and 3 without it. In either case, the interrogative is adverbial. Only in 1937 does WhSV with the present of cleft outnumber WhVS, and both patterns are in complementary distribution. The examples in (3), both attested in the same play, instantiate that:

- (3) a. Mas que **veio você fazer** aqui novamente? (1937 - BP)
 but what came you to-to here again
 ‘But what have you come to do here again?’
- b. Que **é que você veio fazer** aqui novamente? (1937 - BP)
 what is that you came to-do here again
 ‘What are you doing here again?’

- c. E o que **você quer?** (1975 - BP)
and what you want
'And what do you want?'

In the same period, two new structures enter the system: reduced cleft-structures (without the copula) (6a) and Wh-*in-situ* sentences (6b):²

- (6) a. Se eu estiver mesmo grávida, o que **que eu vou fazer?** (1992 - BP)
if I am really pregnant, what that I go to-do
'If I am really pregnant, what am I going to do?'
- b. **Você vai botar** o quê hoje? (1975 - BP)
you go to-wear what today
'What are you going to wear today?'

Pinheiro and Marins (2012) analyze the diachrony of WhV interrogatives. Their results suggest a parallel change involving the decrease of WhVS order and the decrease of null subjects (see Fig. 2).

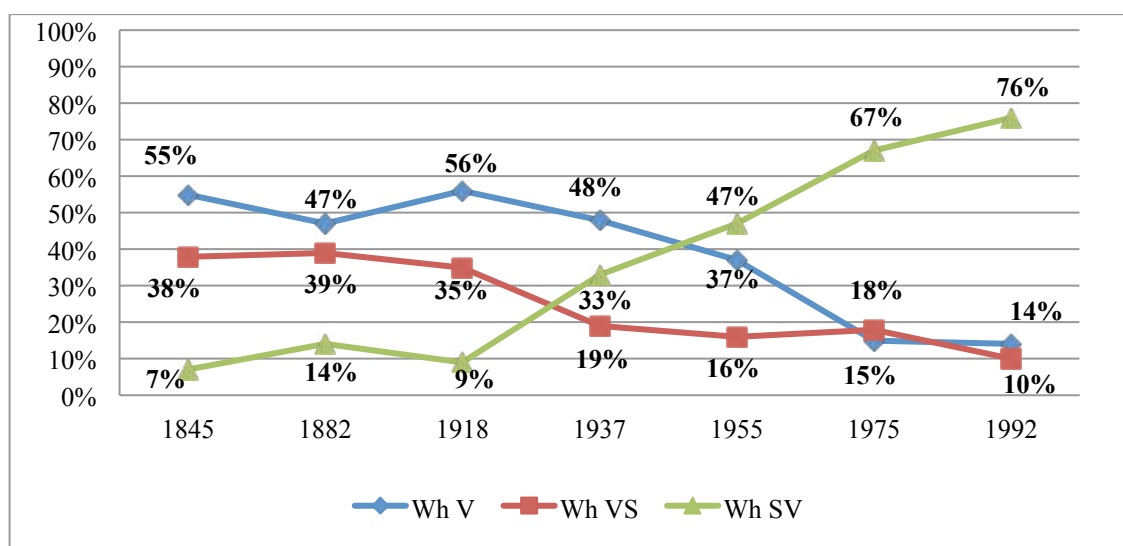


Fig. 2. WhV, WhVS and WhSV sentences in BP (Adapted from Pinheiro and Marins, 2012: 172)

² The author of this article carried out the research for both patterns in the sample analyzed by Duarte (1992).

In Fig. 2, WhSV word order (with and without cleft) starts with a relative frequency of 7% and rises steadily to reach 76% in the final synchrony. Notice that at the same time, the usage frequencies of null subject (WhV) and WhVS constructions descend, reaching almost equal percentage rates in the final period analyzed. Null 1st, 2nd and 3rd person subjects in *Wh*-interrogatives (7) are predominant in the first four periods of time, ranging from 55% to 48%):

- (7) a. Com quem **tenho** o prazer de falar? (1845 - BP)
 with whom have.1SG the pleasure to speak
 ‘With whom do I have the pleasure to speak?’
- b. Para que **estudaste** tanto, rapaz? (1882 - BP)
 for what studied.2SG so-hard, boy
 ‘Why have you studied so hard, boy?’
- c. Onde se **esconderia** ? (1845 - BP)
 where SE.Cl.Refl would-hide.3SG
 ‘Where would she hide?’

To sum up: null subjects are predominant throughout the 19th century and the first half of the 20th century. When an overt subject is used, VS is preferred to SV order in the three first synchronies; from then on, the usage frequency of SV order begins to rise and propagates quickly, particularly from the 1950s on. As shown above, the remaining cases of WhVS attested in the last quarter of the 20th century can no longer be considered real “inversions” since they are restricted to unaccusative verbs with a lexical DP subject, an internal argument in fact still allowed in declaratives with such verbs. This is also the moment when two innovations are attested in the analyzed sample: the reduced cleft and interrogatives with *Wh-in-situ*.

2. Theoretical questions and methodology

2.1. *Wh*-questions in European Portuguese

Formal descriptions about the order in EP *Wh*-questions usually start with the contrast exhibited in (8). According to Ambar (1987; 1992), a sentence like (8a) is ungrammatical whereas (8b) is grammatical. This implies that movement of V to C is obligatory in EP, unless the cleft structure is present (8c):

- (8) a. *O que **a Maria comprou**?
 what the Maria bought
- b. O que **comprou a Maria**? (EP)
 what bought the Maria
- c. O que *é que* **a Maria comprou**? (EP)
 what bought the Maria
 ‘What did Maria buy?’

The author’s hypothesis for the obligatory movement is that there is an empty NP category in the *Wh*-structure; therefore, the raising of the verb would fill this empty slot. With complex *Wh*-structures (*Wh* + NP), the order is optional:

- (9) a. Que vinho **bebe o João** habitualmente? (EP)
 what wine drink.3SG the João usually
- b. Que vinho **o João bebe** habitualmente? (EP)
 what wine the João drink.3SG usually
 ‘What wine does João usually drink?’

A more recent study by Brito, Duarte and Matos (2003: 472) confirms Ambar’s analysis, and adds that not only SV but VS can occur with cleft (10b):³

- (10) a. Onde *é que* **a Maria trabalha**? (EP)
 where is that the Maria work
- b. Onde *é que* **trabalha a Maria**?
 where is that work the Maria
 ‘Where does Maria work?’

Contrary to what we have seen for BP, however, Brito, Duarte and Matos (2003) show that the reduced cleft structure is not grammatical in EP:

³ Miotto and Lobo (2016) maintain the ungrammaticality of WhSV without the cleft structure.

- (11) *Quem *que* chegou?
Who that arrive?
 ‘Who arrived?’

As for the occurrence of *Wh-in-situ*, according to the authors, they can appear in ordinary *Wh*- questions or in echo questions, just like in BP, and the interpretation will depend on the intonation.

- (12) a. **A Maria sai** quando? / A Maria sai QUANDO? (EP)
 the Maria leaves when
 ‘When does Maria leave?’
 b. **Ela demorou** tanto por quê? / Ela demorou tanto POR QUÊ?
 she took so-long why?
 ‘Why did she take so long?’

The descriptions by Ambar and Brito, Duarte, Matos above suggest VS is the usual pattern and WhSV with cleft is a less frequent and more constrained order. None of the descriptions reviewed here are concerned with frequency of use or with spoken language nor should they be. This analysis will investigate not only the preferred patterns but their distribution over time EP popular theater plays, which, as shown in a number of studies (cf. Duarte, 2012), allow to conclude that they approach speech.

In addition to the descriptions offered for EP, the theoretical support for this research comes from the Principles and Parameters framework. The central hypothesis is based on the results found for BP, shown in section 2. Since EP is a consistent null subject language (HOLMBERG, 2010; ROBERTS and HOLMBERG, 2010) and does not seem to be undergoing any changes involving the setting of the Null Subject Parameter, I expect null subjects to be very productive, particularly when deictic (1st and 2nd persons) or anaphoric; overt subjects will appear preferably in WhVS pattern and the occurrence of WhSV will be, according to descriptions, constrained by clefting. This situation would reflect the results depicted by Pinheiro and Marins (2012) for BP in the 19th century and the first half of the 20th century, in Figure 2. In other words, I expect the use of *Wh*-questions in contemporary EP to pattern older stages of BP.

2.2. Methodology

I collected data from Portuguese plays produced across the 19th and the 20th centuries, distributed in seven periods of time, following the periodization proposed by Duarte (1992). The comparison is based on Duarte's (1992) and Pinheiro and Marins's (2012) results for BP.

The sample comprises 40 plays, some in recognized publications and others collected in the library of the University of Lisbon. The distribution over time is not regular, since it follows the periodization of Duarte's sample of Brazilian plays. The 19th Century comprises two periods and the 20th Century, five. The results for several investigations show considerable regularity in the 19th century, and in the first quarter of the 20th century; only from the 1930s on, can we attest changes in BP pronominal system and consequent by-products or evidence of the embedding of such major change. More than one author had to be included to represent each period of time in order to obtain a balanced amount of data per synchrony.⁴ The plays are listed in Appendix 1.

The data selection and processing was based on variationist methodology (TAGLIAMONTE, 2006; GUY AND ZILLES, 2007). Besides the parameter of word order in overt subject interrogatives (WhVS/WhSV), *Wh*-questions with null subjects were included in order to test my predictions. In addition, I controlled for other structural factors, such as the type of Wh constituent, the presence of cleft-structure, the grammatical person of the subject, the use of a full DP or a pronoun for 3rd person over subjects, and the transitivity of the verb.

3. Results

3.1. *Wh*-questions in European Portuguese

A total of 681 data were collected, with the following distribution: 53% WhV, 38% and 9% WhSV (cases of Wh *in situ* were treated separately). The overall distribution confirms the general hypothesis that WhV would be the most frequent pattern,

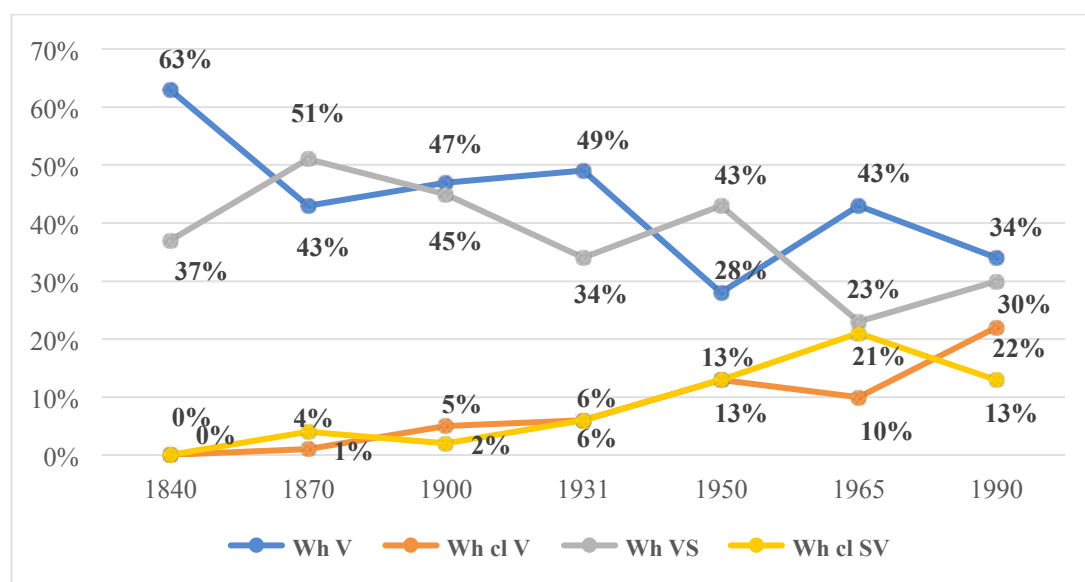
⁴ A research in course investigates individual playwrights who produced in two, sometimes, three periods considered in our sample, so as to observe their behavior across time.

exemplified in (13 a, b), without and with cleft, while (13 c, d) present WhVS and WhSV with cleft, respectively:

- (13) a. Então por que a **agrediu?** (1905 - EP)
 then why her.CL.3SG hurt.3SG
 ‘Then why did he hurt her?’
- b. O que *é que* **quereis dizer** no tal programa? (1871 - EP)
 what is that want.2PL to-say in-the such program
 ‘What do you want to say in that program?’
- c. Então que **respondeu Sofia** à minha carta? (1843- EP)
 then what answered Sophia to-the my letter
 ‘Then, what was Shophie`s answer to my letter?’
- d. O que *foi que* **tu disseste?** (1954 - EP)
 what was that you said?
 What did you say?

The distribution of the patterns attested along the periods, in Table 1 and Figure 3, provides some interesting results:

| Patterns | 1840-1858 | | 1870-1898 | | 1900-1928 | | 1931-1944 | | 1950-1963 | | 1965-1988 | | 1990-1998 | |
|---------------------|-----------|------------|-----------|------------|-----------|------------|------------|------------|------------|------------|-----------|------------|-----------|------------|
| | N° | % | N° | % | N° | % | N° | % | N° | % | N° | % | N° | % |
| Wh V | 57 | 63 | 39 | 43 | 44 | 47 | 49 | 49 | 31 | 28 | 41 | 42 | 31 | 34 |
| Wh cleft V | - | | 1 | 1 | 5 | 5 | 6 | 6 | 15 | 13 | 10 | 10 | 20 | 22 |
| Wh V S | 34 | 37 | 48 | 51 | 42 | 45 | 34 | 34 | 48 | 43 | 23 | 23 | 27 | 30 |
| Wh cleft V S | - | | - | | - | | - | | - | | 1 | 1 | - | |
| Wh S V | - | | 1 | 1 | - | | 4 | 4 | - | | - | | - | |
| Wh cleft S V | - | | 4 | 4 | 2 | 2 | 6 | 6 | 14 | 13 | 21 | 21 | 12 | 13 |
| Wh-in-situ | - | | 1 | 1 | 1 | 1 | 1 | 1 | 4 | 3 | 3 | 3 | 1 | 1 |
| Total | 91 | 100 | 94 | 100 | 94 | 100 | 100 | 100 | 112 | 100 | 99 | 100 | 91 | 100 |

Table 1. Distribution of *Wh*-question patterns across seven synchroniesFig.3. Patterns of *Wh*-questions in EP plays across seven synchronies

The four main patterns in Table 1, represented in Figure 3, show that the null subject (*WhV* pattern) and the overt subject in *WhVS* pattern are predominant, even though the *WhV* pattern declines over time, starting with 63% in the first synchrony to reach 34% in the years 1990, the final synchrony in our sample. The *WhVS* pattern is

more stable across time. The introduction of cleft structures in the second synchrony, just as in BP, triggers not only Wh-*cl*-SV order, but appears in Wh-*cl*-V sentences (with a null subject) as well. Both patterns start with only 4% and 1% respectively and maintain low usage frequencies until the 1930s. From then on while VS and null subjects keep more regular frequencies. A closer examination of the three last time frames, with the presence of cleft in two patterns, announces what the results for speech recorded in two time frames for EP and BP will show. Its increasing use, either with and over subject in patten Wh SV and with a null subject may reinforce the role of clefting as a multi-functional strategy to codify focus in Portuguese, as pointed out by Kato and Martins (2016).

Only one sentence was attested with VS order and the presence of cleft in the sample analyzed:

- (14) Onde *é que está isso* na comédia? (1957 - EP)
 where is that is this in-the comedy
 ‘Where is all that in the comedy?’

Another structure in Table 1, Wh SV, without cleft, not represented in Figure 3, shows only five occurrences, one attested in the second synchrony (ex. 13a) and four in second quarter of the 20th century (fourth synchrony), the latter produced by the same author, one of them with a complex *Wh*-structure (ex. (13,b,c):

- (15) a. Quem **nós temos** por aí apresentável? (1874 – EP)
 Who we have around here presentable
 ‘Who do we have around here presentable enough?’
- b. Porque **ela** te **horroriza**? (1944 - EP)
 why she you.CL.2SG frighten.3SG
 ‘Why does she frighten you?’
- c. Que noção exata **você tem** do que seja liberdade? (1944 - EP)
 what notion exact you have of-the what is freedom
 ‘What do you know about the meaning of freedom?’

Finally, the last line in Table 1 illustrates *Wh-in-situ* questions, which also appear in the sample from the second period on, keeping very low rates along the time frames (only eleven occurrences attested):

- (16) a. Acusas-me, então **de quê?** (1973 - EP)
 Accuse.2SG-me.CL.1SG then of what
 ‘Of what do you accuse me then?’
- b. E isso é pra **quando?** (1998 - EP)
 and this is for when
 ‘And when is this for?’

3.2. Comparing BP and EP in the writing of theater plays

Figures 2 and 3 in the previous section make it clear that EP and BP show very distinct behavior with respect to the patterns attested for *Wh*-questions. EP is very consistent in the preference for null subjects and VS order, even though we have attested some change in the second half of 20th century. BP, on the other hand, has become a *WhSV* system. Both grammars have in common the introduction of cleft structures, but their roles seem to be different in each system. In EP, as we have seen, it is on the way to become a marker of *Wh*-questions, either with null or overt subjects. Figure 4 shows the course of clefts in EP.

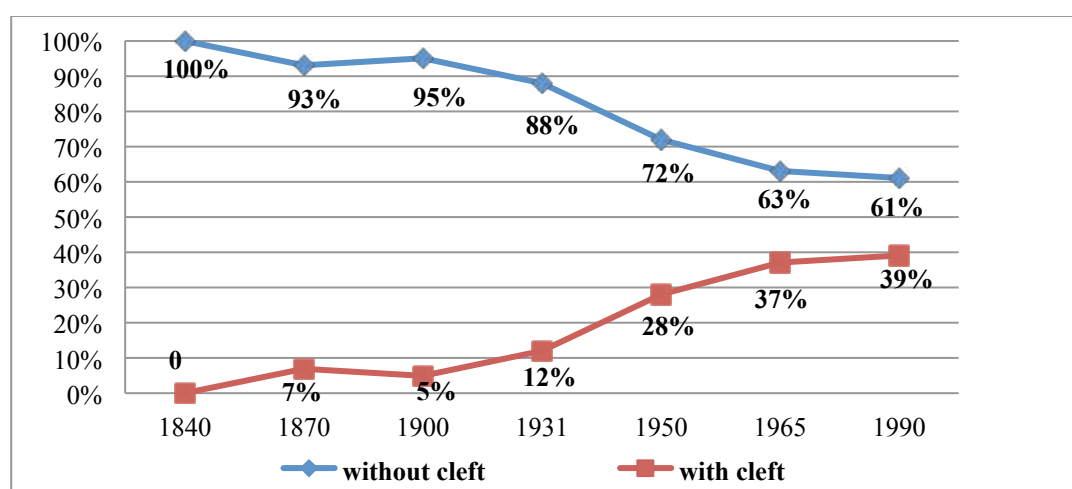


Fig. 4. The evolution of clefts in EP

Questions without clefts are still more frequent in the sample analyzed for EP, which is in part related to the significant occurrence of WhVS order, but cleft questions do show a steady rising trajectory. As for BP, cleft structures have had an important role in the implementation of SV order, and, since null subjects and VS order have almost disappeared, they can be considered the canonical type of *Wh*-questions, even though WhSV can still occur without cleft. Figure 5 shows this evolution in BP, with results from Duarte's (1992) analysis, described in the second section of this article.

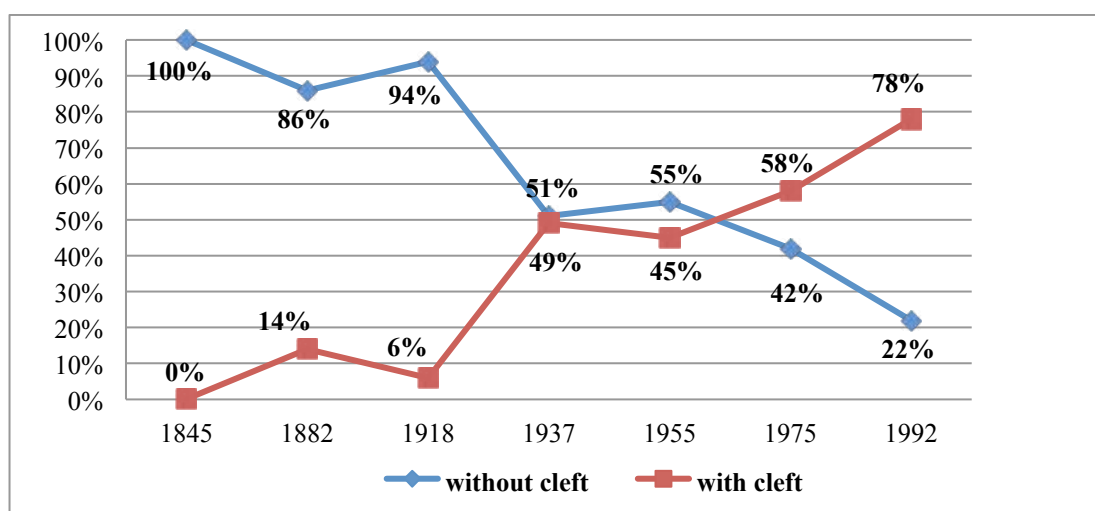


Fig. 5. The evolution of clefts in BP

The propagation of clefting in BP reaches higher rates than in EP, which would suggest that it is a component in focalization in both systems and this feature is not related to the Null Subject Parameter. In fact, based on Kato and Ribeiro's (2009) analysis, Kato (2014) raises the hypothesis according to which *Wh*-questions and focus-structures would have the same checking head (FocusP). Therefore, they should undergo parallel changes across time. This is what the author observes with respect to focalization with VS and with cleft-structures. The canonical cleft structure (copula and complementizer) appears in the 18th century in *Wh*- questions which allows Kato's (2014) suggestion that this innovation will be the preferred focalization pattern from the 19th century on. The results presented here support her hypothesis.

3.3. *Wh*-questions in Brazilian and European speech

In order to compare written to spoken language use, I additionally carried out an investigation of *Wh*-questions in speech. Since in sociolinguistic interviews the

interviewed participants rarely ask questions, I analyzed the interviewers questions. Two samples for European and Brazilian Portuguese recorded in the 1970s and 1980s and in 2010 were used. EP is represented by the interviews belonging to Cordial-Sin and Concordância Projects; BP is represented by Nurc-RJ and Concordância Project⁵. Since the interviewers were teachers or students involved in the research, we can say that the *Wh*-questions collected have been produced by graduates. The results can be seen in Table 2:

| PATTERNS | EP | | BP | |
|--------------------|----------------------|-------------------|------------------|------------------|
| | Cordial-Sin 1980s | Concord. 2010s | NURC-RJ 1970S | Concor. 2010s |
| Wh cleft V | 42% | 53% | 2% | 0% |
| Wh cleft SV | 20% | 16% | 73% | 44% |
| Wh cleft VS | 19% | 9% | 4% | 0% |
| Wh-in-situ | 17% | 5% | 4% | 10% |
| Wh V | 0,5% | 6% | 0% | 3% |
| Wh VS | 1,5% | 10% | 1% | 1% |
| Wh SV | 0% | 1% | 16% | 42% |
| TOTAL (%) | 100 | 100 | 100 | 100 |
| TOKENS | 157 | 120 | 83 | 77 |

Table 2. *Wh*-question patterns in spoken EP and BP

Regarding EP, I first observe that in both synchronies, the interviewers used much more cleft questions than other question types. Regarding subject expression, null subjects are by far the most frequent pattern followed by SV and VS, still significant particularly in the data of the 1980s. *Wh-in-situ* follows, but its use decreases in the second synchrony. The other patterns without cleft are less significant, with the exception of WhV and WhVS in the 2010s. These results confirm the observation about the propagation of clefts in the course of change suggested by Figure 4. I give some examples from spoken EP in (17):

⁵ The Cordial-Sin (*Corpus dialetal para o estudo da sintaxe*) investigates the dialectal syntactic variation in European Portuguese. Available at: <http://www.clul.ul.pt/pt/recursos>. The NURC-RJ Project (*Projeto da Norma Urbana Oral Culta do Rio de Janeiro*) is part of a national project that recorded a large sample of college educated speakers in the five largest capitals in Brazil, including Rio de Janeiro. The interviews were collected in the 1970s and the 1990s. For our analysis the 1970s interviews have been used. They are available at: <http://www.letas.ufrj.br/nurc-rj>. Finally, the data to represent the second synchrony – the 2010s – come from a bi-national project called Concordância, whose purpose was to compare the speech of Rio de Janeiro and Lisbon. The samples, stratified for age, gender and level of education, are available at: <http://www.concordancia.letas.ufrj.br>.

- (17) a. O que *é que* **começaste** por fazer? (Concordância - EP)
 what is that started.2SG for to-do
 ‘What have you started doing?’
- b. O que *é que* **ele ensinava**? (Concordância - EP)
 what is that he taught
 ‘What did he teach?’
- c. E como *é que* se **chama essa operação de ir lá tirar o mel**? (Cordial-Sin EP)
 and how is that se.CL.call this procedure of to-go there to-take the honey
 ‘And how is this procedure of collecting honey called?’
- d. Como **surgiu o gosto pela advocacia**? (Concordância - EP)
 how appeared the vocation for-the advocate
 ‘How did your vocation for law started?’ (When did you decide to be a lawyer?)
- e. E **a peneira é** para quê? (Cordial-Sin - EP)
 and the sieve is for what
 ‘And what is the sieve for?’

As for BP, the distribution confirms Duarte’s (1992) and Pinheiro and Marins’s (2012) findings regarding *Wh*-questions in BP. Null subjects are practically absent and WhSV is the normal word order both with (73% and 44% in each synchrony) and without clefts (16% and 42%). Other patterns are irrelevant, except for 10% of *Wh-in-situ* attested in the second synchrony. I give some examples in (18):

- (18) a. Como *é que* **you escolheu** uma loja de artesanato? (Concordância - BP)
 how is that you chose a store of artcraft
 ‘Why have you chosen to work with artcraft?’
- b. Quando **you vem**? (NURC-RJ - BP)
 when you come
 ‘When are you coming?’
- c. **Vocês viveram** aonde? (Concordância - BP)
 you lived where
 ‘Where did you live?’

Another feature characterizing Brazilian Wh-questions is the existence of reduced clefts (i. e. without the copula), absent in EP:

- (19) a. E *que que você acha* da política atual? (Concordância - BP)
and what that you think of-the politics nowadays
'And what do you think about our politics today?'
- b. Por *que que você acha* que tem esses problemas lá? (Concordância - BP)
for what that you think that have.3SG these problems there
'Why do you think such problems exist there?' (have = there are)

Final remarks

The analyses and comparisons with previous research presented in this paper allows the following conclusions:

(a) EP *Wh*-questions show three main patterns: WhV is the most frequent pattern, obviously with 1st and 2nd person and anaphoric subjects, and expected behavior from a consistent Null Subject Language; the introduction of cleft-structures is decisive in the implementation of WhSV order, but it expands to all patterns, including WhV and WhVS. Therefore, SV order in Wh-questions cannot be associated to the Null Subject Parameter in EP.

(b) WhVS order in BP, on the other hand, is concomitant to the loss of null subjects (Duarte, 1992; 1993; Pinheiro and Marins, 2012). The introduction of the cleft-structure, triggered the change towards WhSV, as shown previously; although still frequent, and usually "reduced" in speech (i. e. without the copula), clefting is no longer a constraint to SV order in BP. As WhSV order increases, *Wh-in-situ* also becomes an important strategy, particularly in the last quarter of the 20th century.

(c) The comparison of the use of *Wh*-questions in recent popular plays to their use in the spoken language of interviewers from Brazil and Portugal has revealed that the use of *Wh*-questions in plays reflects is comparable to their use in spoken language. In addition, the speech data reveals even more clearly the propagation of cleft usage to all patterns of *Wh*-questions in EP, which seems to be an important element to mark focalization (cf. Kato and Ribeiro (2009) and Kato (2014)).

(d) Cleft is also an important structure to focalize Wh constituents in BP, but speech shows a regular distribution between its presence and absence. The most important difference between EP and BP is the robust occurrence of null subjects and VS,

followed by SV – preferably with cleft - in the former, and the absence of WhV, the predominance of SV in the latter, with or without cleft; VS order in BP resists with monoargumental (unaccusatives and copula) verbs and lexical subjects in variation with SV.

Appendix 1. Portuguese plays used in the analysis

| Period I (1841-1857) | Authors (year of birth) |
|-------------------------------------|-------------------------------------|
| Os logros numa hospedaria (1841) | Paulo Midosi (1790) |
| Uma cena de nossos dias (1843) | Paulo Midosi (1790) |
| Casar ou meter freira (1848) | Antonio Pedro L. de Mendonça (1826) |
| Nem tudo que reluz é ouro (1849) | João de Andrade Corvo (1824) |
| O misantropo (1852) | Paulo Midosi (1790) |
| A domadora das feras (1857) | Luís Augusto Palmeirim (1825) |
| Period II (1870-1897) | Authors |
| A liberdade eleitoral (1870) | Teixeira de Vasconcelos (1816) |
| Clero, Nobreza e Povo (1871) | César de Lacerda (1829) |
| Quem desdenha... (1874) | Pinheiro Chagas (1842) |
| O festim de Baltazar (1892) | Gervásio Lobato (1850) |
| A Senhora Ministra (1897) | Eduardo Schwalbach (1860) |
| Period III (1900-1923) | Authors |
| A festa da atriz (1903) | Jorge Santos |
| Terra Mater (1904) | Augusto de Lacerda (1864) |
| Os que furam (1905) | Emídio Garcia (1838) |
| Os Degenerados (1905) | Mário Gollen |
| O álcool (1912) | Bento Mântua (1878) |
| Cavalheiro respeitável (1914) | André Brun (1881) |
| O doido e a morte (1923) | Raul Brandão (1867) |
| Period IV (1931-1944) | Authors |
| Continuação de comédia (1931) | João Pedro de Andrade (1902) |
| Três gerações (1931) | Ramada Curto (1886) |
| A prima Tança (1934) | Alice Ogando (1900) |
| A invenção do guarda chuva (1944) | Luís Francisco Rebello (1924) |
| O ausente (1944) | Joaquim Paços d'Arcos (1908) |
| Period V (1954-1957) | Authors |
| Alguém terá que morrer (1954) | Luís Francisco Rebello (1924) |
| É urgente o amor (1957) | Luís Francisco Rebello (1924) |
| Mário o eu próprio – o outro (1957) | José Régio (1901) |
| O meu caso (1957) | José Régio (1901) |
| Period VI | Authors |
| A guerra santa (1965) | Luís de Sttau Monteiro (1926) |
| A menina feia (1970) | Manuel Frederico Pressler (1907) |
| A sogra (1973) | Alice Ogando (1900) |
| Prólogo Alentejano (1975) | Luís Francisco Rebello (1924) |
| A lei é a lei (1977) | Luís Francisco Rebello (1924) |
| O grande mágico (1979) | Luís Francisco Rebello (1924) |

| | |
|---------------------------------|--------------------------|
| Grito no outono (1980) | Romeu Correia (1917) |
| Period VII (1996-1998) | Authors |
| Um filho (1996) | Luísa Costa Gomes (1954) |
| Quinze minutos de glória (1998) | Jaime Rocha (1949) |
| O céu de Sacadura (1998) | Luísa Costa Gomes (1954) |

Portuguese Plays

MIDOSI, Paulo. Uma cena de nossos dias. In: REBELLO, Luis Francisco. *Teatro português em um acto: volume I 1800-1899*. Editora: Imprensa Nacional da Casa da Moeda, 2003. Escrita em 1843.

MIDOSI, Paulo. Os logros numa hospedaria. Biblioteca do Museu Nacional do Teatro. Escrita em 1841.

MIDOSI, Paulo. O misantropo. Biblioteca do Museu Nacional do Teatro. Escrita em 1852.

MENDONÇA, Antônio Pedro L. Casar ou meter freira. In: REBELLO, Luis Francisco. *Teatro português em um acto: volume I 1800-1899*. Editora: Imprensa Nacional da Casa da Moeda, 2003. Escrita em 1848.

CORVO, João de Andrade. Nem tudo que reluz é ouro. In: REBELLO, Luis Francisco. *Teatro português em um acto: volume I 1800-1899*. Editora: Imprensa Nacional da Casa da Moeda, 2003. Escrita em 1849.

PALMEIRIM, Luis Augusto. A domadora de feras. In: REBELLO, Luis Francisco. *Teatro português em um acto: volume I 1800-1899*. Editora: Imprensa Nacional da Casa da Moeda, 2003. Escrita em 1857.

VASCONCELOS, Teixeira de. A liberdade eleitoral. In: REBELLO, Luis Francisco. *Teatro português em um acto: volume I 1800-1899*. Editora: Imprensa Nacional da Casa da Moeda, 2003. Escrita em 1870.

LACERDA, César de. Clero, Nobreza e povo. In: REBELLO, Luis Francisco. *Teatro português em um acto: volume I 1800-1899*. Editora: Imprensa Nacional da Casa da Moeda, 2003. Escrita em 1871.

CHAGAS, Pinheiro. Quem desdenha... In: REBELLO, Luis Francisco. *Teatro português em um acto: volume I 1800-1899*. Editora: Imprensa Nacional da Casa da Moeda, 2003. Escrita em 1874.

- LOBATO, Gervásio. O festim de Baltazar. In: REBELLO, Luis Francisco. *Teatro português em um acto: volume I 1800-1899*. Editora: Imprensa Nacional da Casa da Moeda, 2003. Escrita em 1892.
- SCHWALBACH, Eduardo. A senhora ministra. Biblioteca da Faculdade de Letras da Universidade de Lisboa. Escrita em 1897.
- SANTOS, Jorge. A festa da atriz. In: REBELLO, Luis Francisco. *Teatro português em um acto: volume II 1900-1945*. Editora: Imprensa Nacional da Casa da Moeda, 2003. Escrita em 1903.
- LACERDA, Augusto de. Terra Mater. In: REBELLO, Luis Francisco. *Teatro português em um acto: volume II: 1900-1945*. Editora: Imprensa Nacional da Casa da Moeda, 2003. Escrita em 1904.
- GARCIA, Emídio. Os que furam. In: REBELLO, Luis Francisco. *Teatro português em um acto: volume II 1900-1945*. Editora: Imprensa Nacional da Casa da Moeda, 2003. Escrita em 1905.
- GOLLEN, Mário. Os degenerados. In: REBELLO, Luis Francisco. *Teatro português em um acto: volume II 1900-1945*. Editora: Imprensa Nacional da Casa da Moeda, 2003. Escrita em 1905.
- MÂNTUA, Bento. O álcool. In: REBELLO, Luis Francisco. *Teatro português em um acto: volume II 1900-1945*. Editora: Imprensa Nacional da Casa da Moeda, 2003. Escrita em 1912.
- BRUN, André. Cavalheiro respeitável. In: REBELLO, Luis Francisco. *Teatro português em um acto: volume II 1900-1945*. Editora: Imprensa Nacional da Casa da Moeda, 2003. Escrita em 1914.
- BRANDÃO, Raul. O doido e a morte. In: REBELLO, Luis Francisco. *Teatro português em um acto: volume II 1900-1945*. Editora: Imprensa Nacional da Casa da Moeda, 2003. Escrita em 1923.
- ANDRADE, João Pedro de. Continuação de comédia. In: REBELLO, Luis Francisco. *Teatro português em um acto: volume II 1900-1945*. Editora: Imprensa Nacional da Casa da Moeda, 2003. Escrita em 1931.
- CURTO, Ramada. Três gerações. In: REBELLO, Luis Francisco. *Teatro português em um acto: volume II 1900-1945*. Editora: Imprensa Nacional da Casa da Moeda, 2003. Escrita em 1931.

- OGANDO, Alice. A prima Tança. In: REBELLO, Luis Francisco. *Teatro português em um acto: volume II 1900-1945*. Editora: Imprensa Nacional da Casa da Moeda, 2003. Escrita em 1934.
- D'ARCOS, Joaquim Passos. O Ausente. Biblioteca da Faculdade de Letras da Universidade de Lisboa. Escrita em 1944.
- REBELLO, Luis Francisco. A invenção do guarda-chuva. In: REBELLO, Luis Francisco. *Todo o Teatro*. Editora: Imprensa Nacional da Casa da Moeda. 1999. Escrita em 1944.
- REBELLO, Luis Francisco. Alguém terá que morrer. In: REBELLO, Luis Francisco. *Todo o Teatro*. Editora: Imprensa Nacional da Casa da Moeda. 1999. Escrita em 1954.
- REBELLO, Luis Francisco. É urgente o amor. In: REBELLO, Luis Francisco. *Todo o Teatro*. Editora: Imprensa Nacional da Casa da Moeda. 1999. Escrita em 1957.
- REBELLO, Luis Francisco. Prólogo Alentejano. In: REBELLO, Luis Francisco. *Todo o Teatro*. Editora: Imprensa Nacional da Casa da Moeda. 1999. Escrita em 1975.
- REBELLO, Luis Francisco. A lei é a lei. In: REBELLO, Luis Francisco. *Todo o Teatro*. Editora: Imprensa Nacional da Casa da Moeda. 1999. Escrita em 1977.
- REBELLO, Luis Francisco. O grande mágico. In: REBELLO, Luis Francisco. *Todo o Teatro*. Editora: Imprensa Nacional da Casa da Moeda. 1999. Escrita em 1979.
- RÉGIO, José. Mário, o eu próprio – o outro. Biblioteca da Faculdade de Letras da Universidade de Lisboa. Escrita em 1957.
- RÉGIO, José. O meu caso. Biblioteca da Faculdade de Letras da Universidade de Lisboa. Escrita em 1957.
- MONTEIRO, Luís de Sttau. A guerra santa. Faculdade de Letras da Universidade de Lisboa. Escrita em 1965.
- OGANDO, Alice. A sogra. Biblioteca do Museu Nacional do Teatro. Escrita em 1973.
- CORREIA, Romeu. Grito no outono. Biblioteca da Faculdade de Letras da Universidade de Lisboa. Escrita em 1980.
- GOMES, Luísa Costa. Um filho. In: GOMES, L. *Duas comédias*. Editora: Relógio D'água. 1996.
- GOMES, Luísa Costa. *O céu de Sacadura*. Editora: Cotovia. 1998.
- ROCHA, Jaime. Quinze minutos de glória. Biblioteca da Faculdade de Letras da Universidade de Lisboa. Escrita em 1998.

References

- AMBAR, Maria Manuela. “Gouvernement et Inversion dans les Interrogatives QU- en Portugais”. *Recherches Linguistiques*, 16, pp. 5-15, 1985
- AMBAR, Maria Manuela. *Para uma sintaxe da inversão verbo-sujeito em Português*. Lisboa: Edições Colibri, 1992.
- BARBOSA, Pilar. *Null Subjects*. PhD Dissertation, Cambridge, Mass: MIT, 1995.
- BRITO, Ana Maria; DUARTE, Inês; MATOS, Gabriela. “Estrutura da frase simples e tipos de frases”. In: MATEUS, M.H.M *et alii*. (Eds.) *Gramática da língua portuguesa*. 5ª ed. rev. e ampl. Coimbra: Caminho, 2003.
- CHOMSKY, Noam. *The Minimalist Program*. Cambridge, Mass: The MIT Press, 1995.
- DUARTE, M. Eugênia L. “A perda da ordem V(erbo) S(ujeito) em interrogativas qu- no português do Brasil”. *D.E.L.T.A. 8 (Especial)*, p. 37-52, 1992.
- DUARTE, M. Eugênia L. “Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil.” In: ROBERTS, I; KATO, M. A. (Eds.) *Português Brasileiro: uma viagem diacrônica*. Campinas: Ed. da UNICAMP, 1993, p. 107-128
- DUARTE, M. Eugênia L. *A perda do princípio “Evite Pronome” no português brasileiro*. PhD Dissertation, UNICAMP, Campinas, 1995.
- DUARTE, M. Eugênia L. (Ed.) *O sujeito em peças de teatro (1833-1992): estudos diacrônicos*. São Paulo: Parábola, 2012.
- DUARTE, M. Eugênia L. and KATO, Mary A. “A diachronic analysis of Brazilian Portuguese wh-questions”. *Santa Barbara Portuguese Studies*, v. VI, p. 326-339, 2002.
- GUY, Gregory R.; ZILLES, Ana M. S. *Sociolingüística Quantitativa*. 1. ed. São Paulo: Parábola Editorial, v. 1, 2007.
- HOLMBERG, Anders. “Null Subject Parameters”. In: BIBERAUER, T. *et alli* (Eds). *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge/ New York: Cambridge University Press, 2010.
- KROCH, Anthony. “Reflexes of grammar in patterns of language change”. *Language Variation and change*, 1, 1989.

- KATO, Mary. “Focus and Wh-questions in Brazilian Portuguese”. In: DION, N. *et alli.* (Eds.) *Linguistic Variation: confronting fact and theory*. Routledge, 2014.
- KATO, Mary and RIBEIRO, Ilza. “Cleft sentences from old to modern Brazilian Portuguese”. In: DUFTER, A. e JACOB, D (Eds.). *Focus and background in Romance languages*. Amsterdam: John Benjamins, 2009, p.123-154.
- KATO, Mary and MARTINS, Ana Maria. “European and Brazilian Portuguese: An Overview on Word Order.” In: WETZELS, L. *et alli.* (Eds.) *The Handbook of Portuguese Linguistics*, Malden, MA: John Wiley & Sons, Inc. 2016, p. 14-40.
- LOBO, Maria. “Sintaxe e Semântica: fenômenos de omissão e eclipse”. In: RAPOSO, E.P *et alli.* (Eds.) *Gramática do Português*, v. II. Coimbra: Fundação C. Gulbenkian, 2013, 2309-2335.
- LOPES ROSSI, Maria Aparecida Garcia. ‘Estudo diacrônico sobre as interrogativas do português do Brasil’. In: ROBERTS, I; KATO, M.A. (Eds.) *Português Brasileiro: uma viagem diacrônica*. Campinas: Ed. da UNICAMP, 1993
- LOPES ROSSI, Maria Aparecida Garcia. *A sintaxe diacrônica das interrogativas Q do português*. PhD Dissertation, Unicamp, São Paulo, 1996.
- NICOLAU DE PAULA, Mayara. *A ordem VS/SV e as interrogativas Q no PB e no PE: uma análise diacrônica*. PhD Dissertation, UFRJ, 2016.
- PINHEIRO, Diogo; MARINS, Juliana. “A trajetória das interrogativas QU- clivadas e não clivadas no Português Brasileiro”. In: DUARTE, M. E. L (Ed.) *O sujeito em peças de teatro (1833-1992): estudos diacrônicos*. São Paulo: Parábola, 2012, p.161-179.
- ROBERTS, Ian; HOLMBERG, Anders. “Introduction: parameters in Minimalist theory”. In BIBERAUER, T, *et alli.* (Eds.) *Parametric Variation: null subjects in Minimalist theory*. Cambridge: Cambridge University Press. 2010, 1-57.
- TAGLIAMONTE, Sali. *Analysing Sociolinguistic Variation*. Cambridge: Cambridge University Press. 2006.