European Portuguese and Brazilian Portuguese: how quickly may languages diverge?

CHARLOTTE GALVES

22.1 Introduction

In 1500, Portuguese-speaking sailors arrived to South American coasts where people spoke Tupi languages. Although in the first part of the colonization period the influence was more from the Tupi speakers towards the Portuguese speakers, as the new incomers learn the language of the natives, the “Terra de Santa Cruz vulgarmente chamada Brasil” will eventually be the greatest Portuguese-speaking language in the world. However, Brazilian Portuguese is in many aspects as different from European Portuguese as from Galician or Spanish. We shall even argue below that from a grammatical point of view, it is much more different, since it displays syntactic features typical of typologically different languages. Portuguese provides us therefore with a case of recent split between two different realizations of a formerly unique language whose effects are observable in the phonetic, lexical and syntactic aspects of the modern languages and can be argued to characterize two radically different languages. In this chapter we shall describe the results and the dynamics of such a split, addressing ultimately the issue of the span of time necessary for languages to diverge. It will be argued that at the core of this question lies the dialectic relationship between I(nternal)-language and E(xternal)-language, in such a way that it is not possible to approach this process without taking into consideration both aspects and the way they interact. This means that we should understand not only the grammatical (I-Language) changes occurred in the history of the new languages and trace back their emergence, but also what provoked them and how they came to be diffused in the population and eventually became sufficiently homogeneous in their output (E-language) to be recognized as a new common language.

With this in mind, the chapter is organized as follows. In Section 2, in order to show that the two languages are effectively different from the point of view of I-Language, the main synchronic syntactic differences between European and Brazilian Portuguese are described, and discussed in the framework of Parameter Theory. Then Section 3 first presents some relevant aspects of the diachrony of European Portuguese,
addressing the question of the language taken to Brazil by the *caravelas*. It is argued that Brazilian Portuguese directly descends from this language and not from Modern European Portuguese, which emerges in the 18th century, although European Portuguese in all its versions never ceased to play a role in the outcome of the change. In the second part of the section, it is shown that the similarity between BP and the Niger-Congo languages pointed out in Section 2 are likely not to be casual but to be due to the intensive contact with African languages created by the importation of millions of African slaves during more than three centuries. Finally, in Section 4, all the previous ingredients are put together in order to propose a model of the dynamics of a change resulting in the split into two different languages from a unique one. As already mentioned above, it is argued that such a change cannot be understood only from an Internal-language point of view, i.e in terms of parameter resetting. It also requires an External-language approach, since it depends on socio-historical processes leading to the diffusion of the new language to the whole population, which involve readjustments between the new and the old language.

### 22.2 European Portuguese and Brazilian Portuguese today: phonological and syntactic differences.

The first question to be addressed is the nature of the differences between European and Brazilian Portuguese today. In this chapter, we shall leave lexicon aside, and concentrate on phonological and syntactic differences.

#### 22.2.1 Phonetic differences

Beyond any punctual phonetic difference between Modern European Portuguese (henceforth EP) and Brazilian Portuguese (henceforth BP), their main distinction is prosodic in nature. It can be said that they evolved in a diametrical opposite way from their common ancestor that we shall call Classical Portuguese (henceforth, ClP). On one hand, it has been shown that, differently from the other Romance languages, EP displays properties of stress-timed languages, mainly expressed by the reduction or complete deletion of non-stressed vowels, both in pre and post-tonic position (references). Phonetically, this creates consonantal clusters that make consonantal length vary a lot, as well as make the number of types of syllables increase. BP, on the contrary, avoids consonant clusters, by the means the epenthesis of /i/, tending to implement a CV alternation reminiscent of moraic
languages. On the other hand, both production and perception studies also evidenced that the two languages have a distinct pattern of attribution of secondary accents to pre-tonic syllables. While it is governed by a binary alternation in BP, it tends to fall on the first syllable of words in EP. Additionally, in EP but not in BP, the attribution of secondary accents is sensitive to the position of the words in the clause (Sandalo et al. 2006; Frota and Vigário 2000, 2001; Reich and Dufter 2003).

Another aspect of the differences in the rhythmic properties of the languages concerns clitic pronouns. We shall come back in the next section to the syntactic aspects of clitic-placement, but it must be mentioned here that BP does not display a property characteristic of European Portuguese along all its history: the prohibition of clitic pronouns in absolute first position in the sentence. With respect to this phenomenon, BP and EP also evolved in a diametrically opposite direction. As we shall briefly describe below, ClP presents a mixed pattern, in which proclisis is largely dominant, except for the fact that enclisis is mandatory in absolute first position, as an effect of the strict application of the Tobler-Mussafia Law. From a descriptive point of view, it is accurate to say that while EP generalized enclisis, BP generalized proclisis. Even if part of the process is syntactic in nature, this opposite evolution reveals a fundamental prosodic difference between the two languages, which is coherent with their respective prosodic patterns: only in EP are first position effects observable both at the phonological word level and at the Intonational Phrase level, favoring enclisis.

22.2.2 Syntactic differences
The main syntactic differences between BP and EP can be divided in the following classes that we shall describe below:

1. Verbal and nominal agreement
2. Pronominal syntax
3. Null categories

Below we will review the properties that strongly differentiate BP and EP syntaxes. In Section 22.2.3, these phenomena will be discussed from the point of view of Parameter Theory.

22.2.2.1 Agreement
Although the variability of verbal agreement has been one of the more noticed property of BP when compared with EP, we shall see below that not only does BP display a variable agreement where it is categorical in EP but it also allows agreement relations where they are completely impossible in EP.

The first aspect has long been acknowledged, and led people to argue that BP underwent a creolization phase (Guy 1981). However, a closer look at the facts shows that agreement is variable in all social classes. Based on several quantitative studies, Lucchesi (2009: 348) shows that its occurrence ranges from 16% in afro-brazilian communities to 73% in educated people from the city of Rio de Janeiro. The amplitude of its variation thus depends on three interacting variables: isolated vs integrated populations, non-urban vs urban context, less education vs. more education.

The variability of verbal agreement also affects person features. As far as person is concerned, it is important to note that in most of the dialects of Brazil the second person of the singular is no longer used in the verbal paradigm. This creates a context for the lack of agreement between the subject and the verb for speakers who use the 2nd person pronoun ‘tu’ instead of, or in variation with, the 3rd person pronoun ‘você’, which in Brazil lost the polite/formal feature that it has in Portugal. It is interesting to note that the use of ‘tu’ is dependent not only on the socio-cultural factors that govern number variation but also on regional factors. In both cases, it induces lack or variability of person agreement.

Less described until recently are the agreement facts correlated with the ‘topic-oriented’ properties of BP, first described by Pontes (1981 [1987]). Pontes showed that BP shares with “Topic-Oriented” languages like Chinese, constructions in which the first constituent DP of the sentence is not the subject but some locative or genitive argument, immediately followed by the verb and the internal argument of the verb:

(1) Esta casa bate muito sol
    this house stike much sun
    ‘There is too much sun on this house’

(2) A Sarinha nasceu o dente
    the Sarinha (was) born the tooth
    ‘Sarinha has got a tooth.’
As discussed recently by many scholars (Galves, 1998; Negrão, 1999; Avelar and Cyrino, 2008; Negrão e Viotti, 2008; Avelar and Galves, 2011, Naves and Pilati, 2012; Andrade and Galves 2014 a.o), there can be an overt agreement between this pre-verbal NP and the verb:

(3) Estas casas batem demasiado sol
these houses stike.3.PL too-much sun
‘There is too much sun on these houses’

The possibility of this agreement strongly distinguishes BP from EP, where it is completely impossible (see the discussion in Costa 2010).

The variability of number agreement observed in clauses in the preceding section is also found in noun phrases, as described by many authors (Guy 1981, Scherre 1988, Holm 2004 a.o), as in the following example (Holm 2004: 101):

(4) um dos mais velho orixás
one of-the.PL more old.SG deities
‘one of the more ancient deities’

22.2.2.2 Pronominal paradigm and syntax

The remarkable facts of the BP pronominal paradigm and syntax are the following:
- First and second person accusative and dative clitic pronouns vary with strong pronouns (5)-(6).
- third person accusative clitic o/a (-him/-her) is substituted in speech by the strong form ele (or by an empty category – cf. 22.2.2.3) (7)
- third person dative clitic ‘lhe’ tends to be used only as a form of address, corresponding to the strong pronoun ‘você’ (you) and consequently ‘lhe’ and ‘te’ freely vary (8). Accordingly, the possessive pronouns ‘teu’ and ‘seu’ freely vary to refer to ‘você’ (9)

(5)a. Eu não te vi ontem
I not CL2 saw yesterday
b. Eu não vi você ontem
I not saw you yesterday
‘I did not see you yesterday’

(6)a. Me leva com você
    CL1 take with you

b. Leva eu com você
take I with you

‘Take me with you’

(7)a.% Eu não o vi
    I not CL3Sg saw

b. Eu não vi ele
    I not saw he

‘I did not see him’

(8)a. Eu te dei esse livro
    I CL 2 gave this book

b. Eu lhe dei esse livro
    I CL3DAT gave this book

‘I gave you this book’

(9)a. Você trouxe teu livro?
you brought POS2 book
Você trouxe seu livro?
you brought POS3 book

‘Did you bring your book?’

Summarizing, all these facts can be interpreted in the following way.

- Clitics referring to the third person of the discourse are no longer licensed by the grammar.

- Consistently with the absence of the 2\textsuperscript{nd} person in the verbal paradigm, the 2\textsuperscript{nd} and the 3\textsuperscript{rd} person are no longer distinct in the pronominal feature system, which suggests that Person in BP is a binary feature (Galves 1993), with the value [+] corresponding to the first person, and the value [-] to the rest.

- Tonic pronouns have no case feature, since they can appear in any position. This claim is supported by the fact that clitic doubling occurs without preposition (10).

(10) Me leva eu contigo
Take me with you

Last but not least, BP differs from EP with respect to clitic-placement. We have mentioned above the fact that clitics can occur in absolute first position, in contrast with what is observed in Portuguese texts from the origins to our days. In accordance, proclisis is the general rule. But it must be emphasized that proclitic placement is not the only difference between EP and BP. They also differ with respect to the category to which the clitic affixes. In the former, it adjoins to the tensed form, be it an auxiliary or the thematic verb. In the latter, it is proclitic to the thematic verb, be it tensed or not. The following examples, from Galves et al. (2005) show the difference between BP and EP exemplified by the revisions performed in the Portuguese edition of Paulo Coelho’s novel ‘O Alquimista’.

(11) a. Me chamo Fátima – disse a moça... (BP)
   CL1 call Fatima said the girl
   b. Chamo-me Fátima – disse a moça (EP)
      call –CL1 Fatima said the girl
   ‘My name is Fatima, said the girl’

(12) a. Como tinha se comportado de maneira correta... (BP)
   As had CL3REFL behaved in manner correct
   b. Como se tinha comportado de maneira correcta...(EP)
      As CL3REFL had behaved in manner correct
   ‘As he had behaved correctly...’

22.2.2.3 Null categories
EP and BP are both null subject and null object languages (Raposo, 1986; Galves 1989; Cyrino, 1997 a.o). However, they differ in the realization of these properties. Null objects are more constrained in EP than in BP, and null subjects are more constrained in BP than in EP, which led many researchers to argue that BP is a partial null subject language (cf. the papers in Kato and Negrão 2000; Holmberg et al. 2009). One of the properties that have been argued to characterize PNSLs is the possibility of interpreting null subjects of tensed clauses as indeterminate. This can be observed in BP, as illustrated in the following dialogue drawn from a comic:
(13)a. Q: Onde está minha cueca de dinossauro?
   Where is my pants of dinossauros
   ‘Where is my pants with dinossauros’
b. A: Está lavando
   is washing
   (i)‘Somebody is washing it’
   (ii)‘It is being washed’

The dialogue in (13) also illustrates the deep relationship existing between indeterminate null subjects and null objects in topic constructions. In fact, the availability of the indeterminate interpretation of the null subject creates an ambiguity concerning the syntactic structure of sentences like (13)b., where two interpretations are available. Either the null subject can be interpreted as indeterminate and the null object as referring to the topic (i), or the null subject can be interpreted as the topic itself, and the value of the sentence is exactly that of a passive (ii). This is still more visible in the following example, drawn from an advertisement, where the topic (‘Tapete de Madeira’ Carpet of wood) is both the thematic subject of the unergative verb ‘encolhe’ (shrink) in b., and the thematic object of the transitive verb ‘encera’ (polish) in c. In the latter, the semantic subject is indeterminate. However the parallelism in the construction suggests that ‘não encera’ can be interpreted as a passive sentence.

(14)a. Tapete de madeira
   carpet of wood
b. Não encolhe
   not shrinks
c. Não encera
   not polishes
   ‘Wood carpet
   It does not shrink
   It does not need to be polished’
These examples show that indeterminate null subjects pave the way for the existence of passive constructions without any passive morphology. This extends to the passive (or indefinite) se-constructions, which, as well as the indeterminate se-constructions tend to occur without se, which is impossible in EP.

(15)a. Aqui conserta sapatos  (BP)
   here repair.3Sg shoes
b. Aqui conserta(m)-se sapatos (EP)
   here repair.3Sg(Pl)-REFL shoes
   Here shoes are repaired

It is therefore fair to say that in the process of change, BP lost morphology: it lost the 2\textsuperscript{nd} / 3\textsuperscript{rd} person distinction, it lost obligatory plural agreement on nouns, and person/number agreement is variable on verbs too. It also tends to avoid passive morphology, third person clitics and se-constructions. This means that BP has less morphology than EP, but differently from Creole languages, it does have flexional morphology. However, its syntax is substantively different not only from EP but also from other Romance languages, since it includes agreement processes typical of other families of languages, as we saw above. This is why, even if it did not go through a creolization process, we can say that Portuguese in Brazil underwent a restructuring process (Holm 2004).

22.2.3 A parametric view of the differences between EP and BP
The differences between BP and EP offer an interesting case study to Parameter Theory, and therefore to Syntactic Change Theory. The review made above evidences that some parameter value differentiates the two languages in spite of the fact that they continue to bear the same name and are, to a great extent mutually intelligible, due to the fact that they share a great part of their lexicon and their phonology and derivational morphology. Following Pontes’ analysis based on typological considerations, many researchers have argued that BP is a topic-oriented language (cf. Galves 1998, 2000; Negrão 1999; Kato and Duarte 2008). This was challenged by Costa (2010) who argued that the locus of the difference between EP and BP was not the possibility of construing topic-comment sentences but the establishment of agreement between the topic and the verb. From that he concludes
that the difference between EP and BP is not macro-parametric, but only micro-parametric.

Costa’s discussion is reminiscent of the conception of the difference between macro and micro-parameter defended by Baker (1996, 2008). In his pioneering discussion about the question, Baker (1996) associated with the notion of macro-parameter what Sapir (1921) had called the “structural genius” of languages. From this point of view, macro and micro-parameters have distinct natures, and only the latter concern the features of functional categories, as proposed in the classic Borer-Chomsky approach. Macro-parameters, by contrast, are the parametrization of the way UG principles are satisfied across languages. Baker (2008:) illustrates his approach with the ‘Agreement Parameter’, which he formulates in the following terms:

(a) A functional head F agrees with NP only if NP asymmetrically c-commands F.

(Yes: Niger Congo languages; No: Indo-European languages)

(b) A head F agrees with NP only if F values the Case feature of NP or vice-versa.

(No: Niger Congo languages; Yes: Indo-European languages)

Baker’s Agreement Parameter is of great relevance for the discussion about Brazilian Portuguese in two related aspects. One of them is that the two large groups of languages it distinguishes are the Niger Congo family and the Indo-European family. This obviously directly concerns the history of Portuguese in Brazil, since there it was acquired and eventually transmitted as a first language by speakers of African, mainly Niger-Congo and in particular Bantu, languages (cf. 22.3), The other aspect is that it can be claimed on the basis of the facts described above that, with respect to Agreement, Brazilian Portuguese functions in great part more like Niger-Congo languages than like Indo-European languages. In effect, agreement can occur with what asymmetrically c-commands the functional features of the verb, independently of its status as subject, as illustrated in (3).

22.3 The roots of BP

22.3.1 The Portuguese roots of BP
One crucial question to be addressed is the nature of the language that was taken to Brazil by the Portuguese colonists. That is, as put in a picturesque way by Castro (1994:138): “há que determinar que tipo de português veio nas caravelas”. It must be noted, however, that the language of the caravelas is not the only one involved since people kept emigrating from Portugal during a long time after the first expedition led by Pedro Alvarenga Cabral. From the second half of the 16th century to the independence of the land, in 1822, Brazil was a destination for hundreds of thousands of Portuguese people, in successive tides. To a great extent, therefore, the comprehension of the formation of Brazilian Portuguese depends on the knowledge of the history of European Portuguese during those three centuries.

At first glance, and according to the traditional periodization of Portuguese (cf. the review in Mattos e Silva 1994), this entails that three different versions of the language went to Brazil: Old Portuguese at the very beginning of the colonization, then Classical Portuguese from c. 1550 to some point of the 18th century, and finally Modern Portuguese. Many authors additionally sub-divide Old Portuguese into Galegan Portuguese, which was the language of the first known written texts, from the second half of the 12th century to around 1350, and Middle Portuguese (Cintra 1963; Castro 1996, 2006) up to the first half of the 16th century. From this point of view, the sailors of the caravelas spoke ‘Middle Portuguese’. Middle Portuguese is characterized by a high rate of variation between old and new forms, at phonological, morphological and syntactic levels (Mattos e Silva 1994, Cardeira 2005, Namiuti 2008 a. o).

Galves et al. 2006 argue that, from a perspective that shifts from E-Language to I-Language, and integrates the notion of Grammar competition (Kroch 1994, 2001), Middle Portuguese should be understood not as the end of the Old Portuguese period but as the beginning of Classical Portuguese, since it is the expression of the emergence of a new grammar in texts. This new grammar enters in competition with Galegan Portuguese, as the latter gradually looses its prestige as a literary language. The new grammar emerges as the result of the displacement of the political and cultural centre of Portugal from the North to the South, which involves a great contact of dialects from all regions of Portugal and the installation of a new urban standard (cf. Silva Neto 1977). This urban standard is also marked by an increasing cultural identification with the notion of Hispania, whose more prestigious linguistic expression is Castilian (cf. Paixão de Sousa 2004). From this point of view, the 16th
century ceases to be a major linguistic boundary. It is just the period in which the new, southern grammar, reaches its full expression in texts.

This entails that the language taken to Brazil by the *caravelas*, is not Old Portuguese, but some early version of Classical Portuguese (henceforth CIP). If spoken language precedes written language in the innovations, we can think that as far as the generation who arrived at the *Terra of Santa Cruz* in the first half of the 16th century is concerned, Galegan Portuguese had but little influence on their speech, except, may be, for people who came from the northern parts of Portugal who may have been kept for a longer period close to the old language. Instead, the language of the majority of the newcomers must have been close to what we observe in the 16th and 17th century texts. Moreover, as argued by Silva Neto (1977), there was likely to be a linguistic levelling of the different dialects imported from Portugal, producing a ‘common linguistic structure’ (Silva Neto, 1977:109). Commenting on the fact that the dialectal phonetic features of the northern dialects are absent from BP, the author argues that this was not due to a greater linguistic influence of the southern dialects, but to the fact that the language from the South of Portugal had already suffered a process of elimination of regionally marked phonetic features during the period of the *Reconquista*, in which people of all the land had entered in contact. Postponing to the next section the question of the role of EP in the formation of BP, we can conclude that of the grammatical basis for BP is likely to have been CIP. In order to bring linguistic evidence for this claim, I shall focus now on two aspects of this grammar, which are particularly salient when we compare it with EP. One concerns clitic placement and the other the position of the verb and the subject.

As mentioned in Section 22.2, in the CIP texts of the 16th and 17th century, clitics never appear in first absolute position in matrix tensed clauses. They are thus categorically enclitic in this context. However, at the same period, when the verb is preceded by some phrase, enclisis is marginal, ranging from 0 to 20%. Galves, et al. (2005)b. argue that CIP is sensitive to the Tobler Mussafia Law, which prevents the appearance of clitic pronouns at the initial position of an Intonational Phrase. According to this view, the cases in which V2 sentences are enclitic derive from the existence of an intonational boundary between the pre-verbal phrase and the verb.

The second aspect is that CIP displays a V-to-C syntax (Galves 1997; Paixão de Sousa 2004; Galves and Paixão de Sousa 2005; Antonelli 2011; Gibrail 2010; Galves and Gibrail 2013; Galves and Paixão de Sousa 2013), as was also argued by Ribeiro
(1995) for Old Portuguese. In both cases, there is evidence that the verb raised to a high position (Comp, or Fin in a cartographic approach, Antonelli, 2011). However, they both differed from canonical V2 languages by the facultative character of the presence of a phrase in pre-verbal position. V1 is in fact a rather frequent order (Cavalcante et al. 2010), both when the subject is null, CIP being a consistent null subject language, and when the subject is realized in post-verbal position.

In the texts written by the generation born at the beginning of the 18th century, we observe changes that signal the emergence of the grammar of Modern European Portuguese (EP), and affect both clitic-placement and the V2 syntax. The more salient of them concerns clitic-placement. In the V2 contexts in which enclisis appeared marginally, this placement becomes more and more frequent, arriving at the 19th century close to the modern categorical post-verbal placement. At the same time, VS order drops abruptly in the space of a generation and the frequency of non-subjects in first position dramatically decreases. This strongly suggests that the V-to-C syntax was lost. Figure 1 shows that the decrease of VS (grey triangles) and the increase of enclisis in SV (black square) coincide in time.

![Figure 1: The evolution of VS order and of enclisis in SV in matrix tensed clauses (16th-19th c.)](image)

**Figure 1**: the evolution of VS order and of enclisis in SV in matrix tensed clauses (16th-19th c.)
It must be noted that the evolution of Portuguese in Brazil was in part similar, since BP is not a V-to-C language, and in part different, since, as was already mentioned, clitic placement in BP is predominantly pre-verbal, even in absolute first position. The first fact could lead us to conclude that BP did not evolve from ClP but from EP. However, in the next section, we shall see that there are facts that suggest that the shared loss of V-to-C is not due to the fact that BP is the daughter of EP, but that they independently evolved from ClP. On the other hand, we shall see that EP did influence BP, but by the means of the pressure of the norm.

It must be emphasized, however, that from a historical point of view, and given the periodization proposed above, the hypothesis that the basis for BP was not ClP but EP cannot be a priori discarded. This is because, during a long time, Portuguese was not the dominant language in all parts of the new colony, as the native population did not massively acquire Portuguese while the Portuguese colonists did learn the native languages. This situation produced tupi-based contact languages known as ‘Linguas gerais’, (litterally, general languages), and only in the 18th century will Portuguese become the dominant language in the whole country, due to the aggressive linguistic policy put forward by the Portuguese government. This is at the time when, in Portugal, the dialect spoken by the dominant classes acquires features that eventually lead to a rhythmically distinct language and a new grammar emerges (see Section 22.2 above). This is also a little after hundred of thousands Portuguese people emigrate to Brazil in the rush for gold. The convergence of those facts could militate in favour of the predominant role played by EP as the basis for BP. There is however linguistic evidence that, at that time, the new dialect emerged in Brazil had already its own configuration and was only indirectly affected by the new contingents of Portuguese colonists.

There are several facts that suggest that EP is not at the basis of BP and that in the 18th century the two languages were already evolving independently. In first place, the prosodic change mentioned above did not affect the language spoken in Brazil. As we saw above, BP and EP strongly contrast in their pronunciation, and this is due to a great extent to the innovations that occurred in Portugal from the end of the 17th century on (Révah 1952). The fact that such innovations were not transmitted to the language spoken in Brazil is an indication that EP was not the basis for BP.
Furthermore, there are some pieces of evidence that in the 18th century, at least some aspects of the pronunciation of Portuguese in Brazil were already different. The Portuguese grammarian Frei Luís do Monte Carmelo mentions in 1767 as a Brazilian feature the non-distinction between open pretonic vowels like in ‘pádeiro’ (baker), ‘prègar’ (preach), ‘còrar’ (colour) and closed pretonic vowels like in cadeira (chair), pregar (ham), morar (live). Moreover, the variation resulting from the double colonization of the South of Brazil in the 18th century by Brazilians from São Paulo and immigrants from Azores also shows that phonetic features like the palatalization of -t- and the use of você, already differentiated the ‘Paulistas’ from the new colonists from the Azores, who kept ‘tu’ and do not have palatalization (Oliveira, 2001).

As far as syntax is concerned, we can divide the relevant facts in two classes. First, as in the case of prosody, we observe that changes that occurred in European Portuguese in the 18th century did not affect BP. Second, in cases in which both EP and BP differ from CIP, it is possible to argue that BP is closer to CIP than to EP, and that therefore, it is more likely that BP emerged as the result of changes that affected CIP than the effect of changes that affected EP. In the first class, I shall consider the use of determiners and the expression of progressive aspect, and in the second class the two phenomena that were already described above: clitic-placement and the syntax of topicalization.

One remarkable change between CIP and EP is that some contexts in which definite determiners were optional in the former require their presence in the latter. The more studied of those cases are possessive DPs in argumental position (Castro 2006, Floripi 2008), in which the presence of a definite article before the possessive pronoun was variable in CIP and became obligatory in EP, as illustrated in (15):

(15)a. o meu amigo CIP/EP/BP
    the my friend

b. meu amigo CIP/*EP/BP
   my friend

Interestingly, BP patterns with CIP, since it also allows variation in the use of definite determiner in this context (Oliveira 1981, Callou and Silva 1997).
Another Portuguese innovation not shared by BP is the use of the periphrasis ‘a + infinitive’ instead of the gerund to express progressive aspect:

(16)a. Os alunos estão lendo CIP/*EP/BP
   the students are reading

b. Os alunos estão a ler *CIP/EP/*BP
   the student are to read
   ‘The students are reading’

This construction is absent from the 16th and 17th century texts of the Tycho Brahe Corpus. It begins to appear in the 18th century, when it enters in competition with the use of the gerundive. The latter disappears in the texts written by authors born in the 19th century. In BP, progressive aspect continues to be expressed by gerundive. In both cases, the null hypothesis is that BP is as it is because it kept the properties found in ClP. If it were an offspring of EP, an extra change would have to occur, to get back to the state anterior to the change leading to EP.

The second class of phenomena is slightly different since both EP and BP are different from ClP. The argument now is that BP is closer to ClP than to EP. As was mentioned above, the syntax of clitics is very different in ClP and in EP. In V>1 order, clitics are proclitic in 80 to 100% of the cases in the texts written in the 16th and 17th centuries (i.e representative of the former), and are enclitic in 80 to 100% of the cases in the texts written in the 19th century (i.e representative of the latter). In BP, as we saw in Section 22.2, clitics are proclitic in all contexts, and in particular in V1. If we hypothesis that BP comes from ClP, the change is straightforwardly accounted for by the claim that the Tobler Mussafia law was lost in BP for prosodic reasons, with the effect of extending proclisis to V1. A change from EP to BP would be much more complex since it would not involve simply an extension of the existing pattern but a complete reversal of the grammar of clitics if we adopt Galves and Sandalo (2012) and Sandalo and Galves (2013)’s analysis that enclisis in EP no longer derives from the Tobler Mussafia law. Both from a point of view of E-language (frequencies) and from a point of view of I-language (grammars), the change would be more unlikely from EP to BP than from ClP to BP.

The last case is trickier, but some elements can be taken into consideration in order to make the same point. As revealed above, it has been argued that CIP was a
V-to-C language. This means that the pre-verbal position was not reserved for subjects but could be used to move topicalized phrases of all kinds.

Paixão de Sousa (2012) argues that this was the basis of the TOP V order found in BP, illustrated by the examples (1)-(3) above, with the difference that in this construction, external arguments are cancelled. In this case, according to the author, the fact that agent subjects were very frequently null in CIP paved the way to the reinterpretation of this kind of construction as non-agentive ones. According to her, “the class of arguments which tend more strongly to surface as null in Classical Portuguese texts is precisely the class of arguments that tend to be suppressed form the verb grid in Brazilian Portuguese: Agentive arguments that function as subjects of broad-valency verbs. “ Thus, probably under the influence of the Niger-Congo languages, the V-to-C XV\textit{pro} was reanalyzed as XV, with possible agreement between X and V. It would be less plausible that such a change were based on a SV language like EP

The claim that BP sprung from CIP does not mean that EP has no effect in the dynamics of change we observe in texts, and in the overall picture of the history of Portuguese in Brazil. In the 19\textsuperscript{th} century, at the same time phenomena that are typical of BP begin to appear in texts, we observe a strong influence of the European norm. This creates a complex picture of grammar competition, in which not only two but three grammars are present, creating a huge syntactic variation. A good illustration of this situation is found in the syntax of clitic-placement, in which both BP and EP compete with CIP, as can be seen in Figure 3, from Carneiro and Galves (2010), based on data from two different corpora: Pagotto (1992) for V2 and Carneiro (2005) for V1 and V2. In this graph, each point corresponds to the rate of enclisis at a given time. We see that at the same time proclitic pronouns begin to show up in V1 (black circles at the top of the graph), in accordance with the grammar of BP, enclisis in V2 (black diamonds and squares) gets more and more frequent, in accordance with the grammar of EP (see Figure 1), until there is a reversal, from 1925 on.

\textbf{Figure 2:} The evolution of the enclisis/proclisis variation in V1 and V2 sentences from 1525 to 1975 (BP)
The influence of EP on BP is important to understand the outcome of the contact between Portuguese and African languages, which will be the topic of the next sub-section.

22.3.2. The non-Portuguese roots of BP

The traffic of African slaves began in 1538 (Mattos e Silva 2004:17) as soon as the colonization of Brazil was fully initiated. According to Mussa (1991:163, APUD Mattos e Silva 2004) in the 17th century, 40% of the inhabitants of Brazil were African and Afro-descendent slaves. And 10% of the population was already half-blood. Furthermore, from the 16th to the 17th century, the proportion of integrated natives drops from 50% to 10%. These numbers express the increasing weight of the Africans in the dynamics of the constitution of the Brazilian population in the first centuries of the colonization. To this demographic situation, stated in static numerical terms, we should add another fact that has a still more important relevance from a linguistic point of view. African and afro-descendent slaves were the main factor of the diffusion of Portuguese throughout the land (Mattos e Silva 2004:82-83; 99-106).xviii

Portuguese was thus, during more than three centuries, massively acquired as a second language, and transmitted as both a second and a first language, by native speakers of African languages and their descendents, who came mainly from Niger-Congo speaking areas, and in majority, from the 17th century on, from Bantu speaking areas (Mussa 1991:145, Bonvini 2008). xix

This indeed accounts for many changes that affected Portuguese in Brazil. It is beyond the aim of this chapter to discuss in detail the features of BP that may be due to the African influence, either directly, through transference of properties, or indirectly, through the effect of general processes of restructuring due to contact/imperfect learning (see Avelar and Galves 2014 and the references therein). We shall only consider some of the characteristics of BP presented in Section 22.2. First, the prosodic tendency to insert vowels in order to avoid consonant clusters is very likely to be the result of the African way of pronouncing Portuguese, since Bantu languages, for instance, are CV(V) languages (Ngunga 2004). With respect to syntax, we saw that BP displays a syntactic construction which, according to Baker (2008) derives from a macro-parameter that strongly distinguishes Niger-Congo languages from Indo-European languages. Another phenomenon is the lack of case-marking on
personal pronouns, which is also a very general feature of Bantu languages. As for the variation in agreement patterns, two mutually compatible hypotheses are plausible. One is that it is due to imperfect learning of morphology, typical of 2nd language acquisition. The other is that it is correlated with the lack of case feature, not only on pronouns but on DPs in general (cf. Avelar and Galves 2013). This would be also due to a transference from Bantu languages, as they do not display morphological reflexes of casual marking.\textsuperscript{xx}

However, it must be stressed that the white Portuguese-speaking population never went below the proportion of 30\% during the colonial period, which leads Lucchesi (2009) to argue that, although million of Africans were brought to Brazil by the traffic during the three next centuries, there were no conditions for the emergence of a creole language in Brazil. The interplay between African languages and Portuguese during three centuries eventually led to what Holm (2004) calls a restructured version of Portuguese, which from the point of view of I-language, corresponds to a new grammar. In the last section, we shall address the issue of the dynamics of the genesis of this new grammar in individuals and its consolidation as a recognized new E-language in the population.

22.4 How quickly may grammars languages diverge: what the dynamics of the emergence of Brazilian Portuguese tells us.

The emergence of a new language implies the convergence and consolidation of many individual evolutions, and their acceptance by the linguistic community. In this sense, it can be said without much doubt that BP appears as a different language during the 19th century (Tarallo, 1993), although its roots are to be found in the preceding three centuries.

When we look at the agents of the birth of BP, we should consider two classes of populations:

1. the afro-descendent population and other L2 speakers’ descendents
2. the L1 speakers’ descendents, both from Portugal (immigrants) and from Brazil (upper classes)

The whole process can be understood as developing through three different phases that are both successive, since they are dependent, and co-occurring, since new L1 and L2 speakers keep entering into the process all the time:
Phase 1: Acquisition of Portuguese as a 2\textsuperscript{nd} language > restructuring 1 (as a L2 language)

Phase 2: Transmission of the restructured language as a 1\textsuperscript{st} language > restructuring 2 (L1 language issued from a L2 language)

Phase 3: Mutual influence between the Brazilian speakers of L1<L2 and the Portuguese and Brazilian speakers of L1<L1

Phase 3b: written versions - Influence of the Portuguese norm.

In this process, we must distinguish the individual level and the populational level. Phase 1 begins to occur at the moment African speakers are in the situation to communicate with Portuguese speakers. This allows us to locate the initial point of the whole process as soon as 1538. It will then take at least a generation for a L1 natural language emerge from the acquisition of this L2 version. This means that the new language, at least some version of it, already existed at the end of the 16\textsuperscript{th} century. But for a long time, it was arguably invisible in the population since it was isolated in members of a non-dominant part of the population. In effect, according to Mussa 1991 (APUD Mattos e Silva 2004), in the 17\textsuperscript{th} century, descendents of Africans correspond to only 20\% of the population, and the way they speak is considered by Europeans as a corrupted Portuguese.\textsuperscript{xii}

Both Phase 1 and Phase 3 are the key to the question of creolization, much debated in the literature on the formation of Brazilian Portuguese, with a great polarization between the hypothesis that the language had undergone a phase of creolization (Guy 1985) and the hypothesis that what had happened was simply a drift (for a recent defence of this idea see Naro and Scherre 2007). In effect, in the absence of any influence of the native and native-descendent speakers of Portuguese on the learners of restructured Portuguese, one of the probable outcomes would be the emergence of a creole language. But, as mentioned above, Portuguese speakers and their white descendents never went below 30\% in the population. Furthermore, the robust presence of native speakers of Portuguese also ensures that its acquisition as a second language by African slaves will not produce in general as drastic effect as what happens in Creole languages. In fact, native speakers played a role at the level of both the acquisition of Portuguese as a second language and its output as the first language of the descendents of the 2nd language learners. Up to now, it has not been possible to prove the existence of creole languages in Brazil and the comparison
between the dialects spoken by contemporaneous isolated afro-descendent communities and Portuguese-based creoles show that the former do not display the properties of the latter (Rougé 2008, Lucchesi et al. 2009).

Phase 3 is also crucial to explain the outcome of Portuguese in Brazil, since the influence was not unilateral. Little by little the language spoken by L1 speakers and their descendents will be transformed by the means of the co-existence and communication with L2 descendents. This process will last during all the colonial period, and give observable fruits at the end of the 18th century, when the different way Brazilian people speak is noticed in grammars and represented in plays (Noll 2008, Ch. ). It is the constant interplay between the L2>L1 native speakers of Portuguese and the L1 and L1>L1 speakers that will ultimately frame the language that becomes to be visible roughly two centuries after it arguably emerged as new I-languages in part of the population speakers’ minds. The 19th century will witness the consolidation of this process, leading, after the independence in 1822, to a nationalist claim of the legitimacy of the new language. But it is only in the 20th century that the more vernacular aspects of it will begin to be acknowledged. At the same time, with the crescent urbanization and education (Lobo 2001), the vernacular continues to be affected by the pressure of the norm, still influenced by the written standard, in particular with respect to the inflectional morphology.

In conclusion, from a E-language point of view, we can say that Brazilian Portuguese today represents a compromise between some aspects of African syntax, and the remainings of a Indo-European flexional system. From a I-language point of view, the challenge is to figure out whether it is generated by only one grammar or it is the outcome of competition of grammars. In the latter case, it is instable, and the history is not over. I leave this issue for further research.

« Land of the Holy Cross, commonly called Brazil » is the title of one of the first description of Brazil, by Pero Magalhães de Gândavo, 1575.
We shall keep ‘EP’ and ‘European Portuguese’ as two distinct entities, with the former referring to the modern language and its grammar, and the latter globally meaning ‘Portuguese in Portugal’ from its first written manifestations at the end of the 12th century up to our days.

For computational or mathematical modelling of the prosodic differences between BP and EP, see Duarte et al. 2002; Sandalo et al. 2006, Galves et al. 2012.

Another remarkable characteristic of BP noun phrases concerns definite determiners. As extensively discussed in the literature (cf. a.o Munn and Schmitt (2001), Schmitt and Munn (2003), and Oliveira and Rothstein (2011)), bare nouns are more frequent and more freely licensed in BP than in EP. In subject position, plural bare nouns can receive generic as well as existential interpretation (in EP only generic interpretation is available). Singular bare nouns are possible in subject and object position with generic interpretation.


‘Você’ is a phonological reduction of ‘Vossa Mercê’, litterally Your Grace, whose use initially reserved for the King (cf. Cintra 1972), lost its strong politeness value, and eventually was substituted for ‘tu’ (you) in most dialects of Brazil.

Although Pontes suggests that there is agreement between the pre-verbal DP and the verb, she does not give examples in which this argument is morphologically visible.

‘One should determine what kind of Portuguese came in the Caravelas.’

Fernão Oliveira (born in 1507 in the northern city of Aveiro), in his Grammatica da Lingoagem portuguesa (1536), displays some features of Old Portuguese like cliticization of the definite determiner onto the preceding verb, the locative clitic hi in the existential locution ‘hi ha’ there is.

The elements of description of Classical Portuguese proposed here are based on the data provided by the Tycho Brahe Corpus of Historical Portuguese (www.tycho.iel.unicamp.br/~tycho/corpus)

A remarkable exception are the Sermons of Padre Antonio Vieira, born in 1608 (Martins 1994), Galves (2002) shows that this is due to the high frequency of contrastive topic in pre-verbal position. It must be noted that, in his letters, Vieira is as proclitic as his contemporaries.

In the 16th and 17th texts, around 50% of the sentences have null subjects.
According to Silva Neto (1977: 72), the only place in Brazil where the Portuguese population was much higher than the native population at the end of the 16th century was Pernambuco, but it was also where the contingent of African was the highest.

In 1758, the Portuguese prime-minister, o Marquês of Pombal, prohibits the use of the ‘língua geral’ and renders obligatory the use of Portuguese in Brazil. (cf. Teyssier, 1980: 95). However, an important Amazonian ‘Lingua geral’, called Nheengatu (‘good language’ in Tupi) survived up to the 20th century.

Mattos e Silva quotes Darcy Ribeiro’s words about ‘Brazilian black people’ “[…] succeeding dominating the new language, not only they transformed it, giving singularity to Portuguese spoken in Brazil, but also made possible its diffusion throughout the whole territory”. (In O povo brasileiro. A formação e o sentido do Brasil.)

Beginning with the pioneer work of Mendonça (1935), many scholars argue that Bantu languages played indeed a determinant role in the formation of BP. Evidence for this claim comes additionally from the similarities found between BP and the versions of Portuguese spoken in Angola and Moçambique (cf. Inverno 2011 for Angola and Gonçalves 2010 for Moçambique). Bonvini (2008:34-39) argues that in the 17th century, Quimbundo was extensively spoken in Brazil as a vehicular language. It must be noted that the first grammar of this language was written in Bahia by the Jesuit Pedro Dias, and published in 1697.

See Diercks (2014) for the claim, also put forth in Avelar and Galves (2013), that the presence of case features is parametrized in languages.

See Noll (2008) for a review of travellers reports.