

# 1 History and Current Setting

---

MARIA TERESA BROCARDO AND CÉLIA  
REGINA DOS SANTOS LOPES

## 1. From Latin to Portuguese—Main Linguistic changes and conditioning factors

The development of Portuguese can be traced back to Latin, whose establishment in the Iberian Peninsula was a consequence of the Roman conquest, initiated in 218 BCE, but completed only about 200 years later. In the early fifth century CE the peninsula was invaded by several Germanic tribes, among which was the Suebi, the only tribe that resisted after the arrival of another group of Germanic invaders, the Visigoths. The Visigoths, who eventually conquered the entire peninsula, were already Latin-speaking before they arrived. Of major importance for the linguistic history of the Ibero-Romance area is the Moorish conquest of the Iberian Peninsula, starting in 711 CE. The northern territories of the peninsula were the refuge of the surviving Christian kingdoms, which would later expand southwards. The varieties of Latin spoken in those areas, where the Arabic influence was naturally more superficial, are at the origin of the Ibero-Romance languages.

Portuguese originally emerged in the northwest of the Iberian Peninsula and later expanded southwards with the Christian Reconquest. This “initial” or “primitive” area of Portuguese roughly comprises what corresponds today to Galiza and part of northern Portugal (see Castro 2006: 64–67). Thus, the designation “Galician-Portuguese” is sometimes used in the literature to refer to this original language unity.

In the following description of the language-formation process, we discuss the linguistic features characteristic of the most ancient periods of Portuguese, signaling the most relevant differences within the Romance context, in particular those that are distinctive to the Ibero-Romance area, for which we take Castilian (Cast.) as our point of reference.

As a Romance language, Portuguese both continues Latin and diverges from it, which means that we can identify features that are diachronically characterizable as more conservative while others are of a more innovative nature. This type of contrast, recurrent in traditional approaches but also considered crucial for diachronic analyses in modern approaches (see Maiden, Smith, and Ledgeway 2011 who oppose “innovation” and “persistence” in their *History of the Romance Languages*), will serve as a guideline for our description.

Several vocalic features of Portuguese evidence tendencies of persistence. Among these features is the continuation of the Latin diphthong /aw/ as /ow/, while more generally in Romance it emerges as a monophthong (*ouro* “gold” < *auru-*, Cast. *oro*). This diphthong was preserved in Portuguese even in unstressed syllables (*ousar* “dare,” *outono* “autumn”).

Exceptions are limited to a few forms with earlier monophthongization (*orelha* “ear” < *aureic(u)la-*, *pobre* “poor” < *paupere-*), sometimes already attested in (Vulgar) Latin. It should be noted that Portuguese /ow/ also originated from secondary \*/aw/, formed through distinct processes (*amou* “(he/she) loved” < *amauit*, *soube* “(he/she) knew” < *sapuit*, *outro* “other” < *alteru-*). Portuguese also preserves /ej/, a diphthong formed in Latin in cases where /j/ became adjacent to /ε/ or /a/ including cases where /j/ derived from a consonant, usually /k/ in coda position, also in contrast with the more innovative Cast. result (*madeira* “wood” < *materia-*, Cast. *madera*, *leite* “milk” < *lacte-*, Cast. *leche*).

The lack of diphthongization of the Latin vowels /ε/ < Ē and /ɔ/ < Ō also contrasts with their Castilian reflexes (*erva* “grass” < *hērba*, Cast. *hierba*, *morte* “death” < *mōrte*, cast. *muerte*) and is different from the other Romance areas, in which diphthongization occurred in several phonological contexts. This feature is particularly relevant for the individualization of Portuguese, since it is indicative of a linguistic borderline between Portuguese and non-Portuguese varieties, including Castilian and some varieties of Leonese (Cintra 1983: 140). The Portuguese vowel system as it appears in stressed syllables is still identical to the one that emerged consequent to the loss of the Latin quantity oppositions, already attested in Latin sources, a system defined as “Proto-Western-Romance” (Vincent 1988: 32): /i/, /e/, /ε/, /a/, /ɔ/, /o/, /u/. Other phonological changes affected consonants, such as the evolution of the medial sequences /jt/, /lj/ and /jl/, where Portuguese preserved the sequence /jt/ and where /lj/ and /jl/ developed into /ʎ/, which contrast with the more innovative Castilian reflexes: *oito* “eight” < *octo*, Cast. *ocho*, *folha* “leaf” < *folia*, Cast. *hoja*, *olho* “eye” < *oc(u)lu-*, Cast. *ojo*.

Conservative tendencies are also manifest at other levels. If we assume that the personal infinitive is the continuation of the Latin imperfect subjunctive (Martins 2001), its formal preservation would represent a particularly conservative feature, since the Latin imperfect subjunctive survives elsewhere only in Sardinian (Vincent 1988: 47). The Latin pluperfect indicative is equally preserved in Portuguese (*mostrara* “(I) had shown” < *monstrara(m)*), as in Castilian, in contrast with the other Romance areas. In past stages of Castilian and Portuguese, pluperfect forms conveyed modal as well as temporal values, depending on the context of use, which later split to become a temporal category in contemporary Portuguese while functioning as a modal category in Spanish. Only later in the history of Portuguese would fully grammaticalized compound forms emerge from periphrastic constructions with *habere* and later *tenere* (imperfect) + past participle (*havia/tinha mostrado*) for the expression of the “past in the past” temporal value. The tmesis of clitic pronouns in future and conditional forms (*mostrar-lhe-ei*, *mostrar-lhe-ia* “(I) will / would show him / her”) is often also considered a conservative feature of Portuguese. Although tmesis was already optional in past stages of the language, it persists in contemporary European Portuguese in more formal styles, whereas in Brazilian Portuguese it is restricted to the written register.

Regarding the Latin inherited lexicon, Portuguese and Castilian share a number of etyma that configure lexical contrasts with other Romance areas: *comer* “eat” < *comedere* (cf. Italian / French *manger* / *mangiare* < *manducare*); *falar* / *hablar* “speak” < *fabulare* (cf. *parler* / *parlare* < *parabolare*); *medo* / *miedo* “fear” < *metu-* (cf. *peur* / *paura* < *pavore*, \**pauura*); *pássaro* / *pájaro* “bird” < \**passaru-* (cf. *oiseau* / *ucello* < *auicellu-*). Forms of this type would indicate a common Hispanic origin (Piel 1989: 12) of a more conservative nature. In other cases, however, the shared lexicon of Portuguese and Castilian is somewhat more innovative: *irmão* / *hermano* “brother” < (*frater*) *germanu-* (cf. French *frère* < *frater*); *cabeça* / *cabeza* “head” < \**capitia* (French *chef* < *caput*); *coração* / *corazón* “heart” < \**coratione-* (cf. French *coeur* < *cor*). The conservative character of some rare examples of Portuguese specific lexical types (*colmo* “straw” < *culmu-*, *adro* “atrium” < *atriu-*, *gume* “cutting edge” < *acumen*) could be explained by the peripherality of the Portuguese area (Piel 1989: 12).

Some traditional interpretative approaches (Baldinger 1972: 104–124) explained conservative tendencies as a consequence of the manner in which linguistic Romanization took place

and the specific location of the Portuguese primitive area, by pointing out that Latin expanded to this area from the south (*Hispania Baetica*), where Romanization had occurred much earlier. The variety of Latin that expanded from the south to the northwest of the Iberian Peninsula was supposedly more “cultivated” and thus linguistically conservative. The peripheral location of the Portuguese primitive area, as well as the fact that it was kept in relative isolation by the Suebi kingdom during the first phase of the German occupation (early fifth to early eighth centuries), would also have contributed to linguistic conservatism (Baldinger 1972: 166, Castro 2006: 58–59).

The decisive phonological changes that set apart Portuguese from the other romance languages—the loss of intervocalic *-n-* and *-l-*—occurred later, after the Arab invasions of the peninsula, and clearly are innovative. The exact dating of the loss of *-n-* remains controversial. Regressive nasalization (*manu-* > /mãno/) must have occurred first (Baldinger 1972: 165, 221) and *-n-* loss (/mãno/ > /mão/ “hand,” with a hiatus in Old Portuguese) may have been initiated during the tenth century (Sampson 1999: 184). Nasality is preserved in many forms (*lana* > *lãa* (*lã* “wool”), *bonu-* > *bõo* (*bom* “good”), *unu-* > *ũu* (*um* “one”)), although in a number of other cases it was later lost (*luna-* > *lũa* (*lua* “moon”), *ponere* > *põer* (*pôr* “put”); *tenere* > *têer* (*ter* “have”)). The loss of *-n-* must be related to the more general process of nasalization, which distinctively characterizes Portuguese in the Romance context (French excepted), and by which oral vowels were nasalized when followed by a syllable-final nasal consonant (*non* > Old Portuguese *n[õ]* (*não* “no”), *grande-* > *gr[ã]de* “big,” verbal endings *-an(t)* and *-un(t)* > Old Portuguese *-[ã]* and *-[õ]*).

During approximately the same period (the tenth century, according to Baldinger 1972 and Teyssier 1982: 15) intervocalic *-l-* was also lost after a hypothetical intermediate stage of velarization: *salire* > \**sa[t]**ir* > *sair* “leave”; *colore* > Old Portuguese *coor* > *cor* “colour.” Intervocalic *-l-* and *-n-* occur, of course, in many non-inherited words (or words derived from Latin stems, cf. *mão* / *manual* “manual,” *dor* “pain” / *doloroso* “painful”) and also in words later reshaped according to the Latin model, e.g., Old Portuguese *mêos* “less” (< *minus*) was later replaced by *menos*, while *pena* and *pêa* “punishment” (< *poena-*) both occur in ancient texts. The consequences of *-n-* and *-l-* deletion are noticeable in the peculiar forms of the Portuguese indefinite and definite articles (and of the accusative personal pronouns, descendants of the same Latin forms). The Latin masculine numeral *unu-* “one” became Old Portuguese /*ũu*/ and coalesced to /*ũ*/ probably by the end of the same period, while the result of the feminine form *una-* > /*ũa*/ still persists in Middle Portuguese and suffered epenthesis of *-m-* only later (written *um*, *uma*, *uns*, *umas* in Modern Portuguese). As for the definite articles, which represent the Latin accusative forms *illu-*, *illa-* (from the demonstrative *ille* “that, near to it / him / her”), the *l-* of the intermediary forms, *lo/a(s)*, was preserved only in contexts of assimilation to another consonant (*pelo* < *per* “by” + *lo*), besides several other forms in past stages (Old Portuguese *todos* < *todos* + *los* “all the,” alternative to *todos os*) and also in pronoun cognate forms (*mostramo-lo* “(we) show it” < *mostramos* + *lo*). Otherwise, *l-* loss was systematic, probably because the article and pronoun occurred frequently in intervocalic position in sentence contexts, thus the forms *o*, *a*, *os*, *as*.

The merger to the affricate /*tʃ*/ of the Latin initial clusters *pl-*, *fl-*, *cl-* is another innovative change relevant to the differentiation of Portuguese. The merger also occurs in Castilian, where it became /*ʎ*/, suggesting a more complex change for the Portuguese result: *plenu-* > Old Portuguese *chêo* > *cheio* “full,” Cast. *lleno*; *flamma* > *chama* “flame,” Cast. *llama*; *clave-* > *chave* “key,” Cast. *llave*. Note that this merger does not occur in other Romance areas (cf. French and Italian *plein/pieno*, *flamme/fiamma*, *clef/chave*). The outcome of this change resulted in a contrast between the new affricate /*tʃ*/ and the fricative /*ʃ*/, originated by a distinct (earlier) palatalization process (roughly Latin /*s*/ influenced by /*j*/), and the testimonies of Old and Middle Portuguese show a systematic distinction—/*tʃ*/ written as <*ch*> and /*ʃ*/ as <*x*>. The two consonants would later merge into /*ʃ*/, except in a dialectally

restricted area of European Portuguese (Cintra 1983: 143). Many doublets evidence distinct diachronic paths (*cheio* / *pleno* “full,” *chamar* / *clamar* “call”), and a few words show a less radical change, manifested only in the evolution of *-l* to *[r]* (*prazer* (verb and noun) “please, pleasure” < *placere*, *fraco* “weak” < *flaccu-*, *cravo* “nail” < *clavu-*).

The outcome of the linguistic changes that originated in the northwest of the Iberian Peninsula expanded toward the south as a consequence of repopulation movements subsequent to the Christian Reconquest of the territories that had remained longer under Arabic influence (Castro 2006: 68–69). This migration was the decisive factor in the making of the European Portuguese linguistic area. Portuguese preserves only a few vestiges, mainly in toponyms, of Romance varieties originally spoken in southern areas, generically referred as *moçárabe* (Mozarabic) (Castro 2006: 62–64). These *moçarabismos* (e.g. *Fontanela*) are clearly identifiable in contrast with (Galician-)Portuguese forms precisely because they preserve intervocalic *-l-* and *-n-*.

## 2. Old and Middle Portuguese

### 2.1. Periodization issues

If we take as a reference the approximate date of the phonological innovations discussed in Section 1, we would propose the tenth century as the period of the “birth” of Portuguese (Castro 2006: 75). Although the relatively late appearance of written records of the Portuguese language, similar to the other Romance areas, results in a considerable gap between the estimated date of the emergence of the language and that of the beginning of its attested history, this intermediate period must not be considered entirely non-attested. It is indirectly attested through extant Latin (or Latin-Portuguese) written texts, the most ancient Latin text of the Portuguese area dating from 882. Currently, there is no absolute consensus about the identification of the most ancient text written in Portuguese. In the mid-twentieth century, the results of the research on this subject undertaken by several scholars (see Martins 1999 for references), had converged to the identification of two texts as the earliest preserved Portuguese written records, the *Testamento de Afonso II* (The Will of King Afonso II), dated 1214, and the *Notícia de Torto*, a undated private document probably written around the same date. However, the results of Martins’ (1999) research have reopened the debate. According to this author, among other manuscripts from the late twelfth and early thirteenth centuries, the *Notícia de Fiadores*, dated 1175, is the oldest Portuguese written text. Because it is a very short document with some superficially Latinate forms and constructions, the discussion about how to characterize its writing, and the writing of other documents from about the same period, as Portuguese, Latin or Latin-Portuguese is not closed. The importance of such texts is, of course, undeniable for the documentation of the complex transitional process from Latin to Portuguese writing.

For the periodization of Portuguese history, different proposals have been put forward, some diverging only in detail, while others differ more significantly, especially regarding the evaluation of the relevant linguistic features and the adequate methodologies. Of major importance in this debate is the type of criteria used for establishing the distinct periods. While one would agree that, at least ideally, the proper periodization should be sustained by both linguistic and external facts, more recent proposals tend to emphasize the relevance of linguistic criteria. Here, we will propose a distinction between two medieval periods (following the unpublished proposal of Lindley Cintra, adopted by many): Old Portuguese, comprising broadly the thirteenth and fourteenth centuries, and Middle Portuguese, dating from the beginning of the fifteenth century to the middle of the sixteenth century. Other proposals consider a broader period, usually referred to as *Português Arcaico* “Archaic Portuguese,” covering the whole of the two periods distinguished above.

The main argument for proposing two distinct medieval periods is that several changes, with relevant linguistic consequences, occurred in the transitional period between them, even though they did not all happen at exactly the same time. In the following section we shall briefly describe these changes, concentrating mainly on their consequences for nominal and verbal inflexion.

## 2.2. Contrasting features and linguistic changes

The analysis of the spelling of written Old Portuguese shows a predominant distinction between nasal endings in *-ã* and *-õ*, consistent with the etymology, as exemplified in Table 1.1.

Texts from the late fourteenth century and in particular from the fifteenth century show that these endings were no longer distinguished. The types of non-etymological spellings include confusion between the two endings (*falõ* “(they) talk,” *razã* “reason”) and, especially in later testimonies within the Middle Portuguese period, “phonetic” spellings (*razão* “reason,” *são* “(I) am” or “(they) are,” *andavão* “(they) were walking,” *ordenarão* “(they) ordered”), which clearly attest the merger of the two endings to a nasal diphthong. The merger also included the sequence *-ão*, from Latin *-anu-* (*mão* “hand,” *irmão* “brother”), corresponding to a hiatus in Old Portuguese, but one which had also evolved into a diphthong as the result of disyllabification of the second element. In the words of Sampson (1999: 195), “The details of the merger of the final vowels [-ã], [-õ] and [-ão] of thirteenth century Portuguese as [-ɐ̃w] and the precise factors which it brought are still uncertain and hence controversial.” A patent phonological consequence of this merger was the loss of uniformity of nasal realizations in different contexts (contrast Old Portuguese *red[õ]do* “round” / *raz[õ]*, *gr[ã]de* “big” / *p[ã]* “bread” with *red[õ]do* / *raz[ɐ̃w]*, *gr[ã]de* / *p[ɐ̃w]* in later stages), but we shall concentrate here on the consequences of this change on other levels. In the nominal inflexion of forms with original *-ã* or *-õ* endings, the merger resulted in increased paradigmatic irregularity, and, thus, synchronically less predictable plurals: *pã* / *pães* > *pão* / *pães* and *razõ* / *razões* > *razão* / *razões* (besides *mão* (<-*manu*) / *mãos*). The change also has consequences in verb inflexion, namely the syncretism of the 3rd person plural of the simple perfect, originally in *-õ*, and the pluperfect, originally in *-ã*. This syncretism may have contributed to the disuse of this particular pluperfect form and maybe of the entire paradigm, at least in non-formal registers, and to the increased frequency of the periphrastic pluperfect. It should be noted

**Table 1.1** Examples of the evolution of Old Portuguese nasal endings *-ã* and *-õ*.

| Old Portuguese                                                                   | Latin            |
|----------------------------------------------------------------------------------|------------------|
| <i>-ã</i>                                                                        |                  |
| <b>Nominal forms</b>                                                             |                  |
| <i>pã</i> “bread,” <i>cã</i> “dog”                                               | -an(e)           |
| <b>Third person plural verb forms</b>                                            |                  |
| <i>amã</i> “love” (present), <i>amavã</i> (imperfect), <i>amarã</i> (pluperfect) | -an(t)           |
| <i>-õ</i>                                                                        |                  |
| <b>Nominal (and other) forms</b>                                                 |                  |
| <i>razõ</i> “reason,” <i>oraçõ</i> “prayer,” <i>nõ</i> “no,” <i>cõ</i> “with”    | -on(e), -on, -um |
| <b>Third person plural verb forms</b>                                            |                  |
| <b>Preterite</b>                                                                 |                  |
| <i>ganharõ</i> “won,” <i>fezerõ</i> “made,” <i>partirõ</i> “left”                | -un(t)           |
| <b>Present of <i>ser</i> “be”</b>                                                |                  |
| <i>sõ</i> “(I) am” or “(they) are”                                               | -um, -un(t)      |

that, although the simple pluperfect is fully preserved in Contemporary Portuguese in more formal registers, it is in practice a defective paradigm, since the 3rd person plural forms are uninterpretable. Unlike other forms (*dissera* “(he/she) had said,” *fizera* “had made,” *saíra* “had left”), the 3rd person plural forms (*disseram*, *fizeram*, *saíram*) are always interpreted as the simple perfect. Note that in the plural counterpart of a sentence like *O ministro reafirmou o que dissera antes* “The minister reaffirmed what he had said before,” *Os ministros reafirmaram o que ?disseram/tinham dito antes*, the form *disseram* would not be interpreted as the pluperfect. The only available pluperfect forms are, in this case, the periphrastic ones (*tinham dito / feito / saído*), even if this is usually not signaled in Portuguese descriptive grammars.

Another feature that distinguishes Old and Middle Portuguese is the predominance, in the earlier period, of past participle endings in *-ud-* (varying in gender / number *-udo / a / s*) for 2nd conjugation verbs (with thematic vowel *e*). In the later period, they appear to be only residual and are eventually entirely replaced by *-id-*, causing a merger with the 3rd conjugation (with thematic vowel *i*) endings, for example *perdudo* > *perdido*, *vençudo* > *vencido*, past participles of *perder* “lose” and *vencer* “win,” like *partido*, past participle of *partir* “leave.” The change is usually described as a result of analogy, triggered by the lack of “structural support” for *-u-* (Câmara 1975: 159) in other forms of the verb class. We may assume that both the inter-paradigmatic factors (the influence of the past participle *-id-* endings of the 3rd conjugation verbs) and the intra-paradigmatic factors (the influence of *-i-* endings that already occurred in the 2nd conjugation verbs in the simple perfect and imperfect (*venci*, *vencia* “(I) won, was winning”) concurred for the loss of the *-ud-* endings in past participles. We would then have a change characterizable as either analogical extension or analogical (paradigmatic) leveling, although there may be an additional argument to support the former. The 2nd conjugation lost to the 3rd conjugation a number of verbs, e.g., *caer* > *cair* “fall,” *finger* >  *fingir* “pretend,” *confonder* > *confundir* “confuse” at approximately the same period (Maia 1986: 726–731). This prevalence of the *i*-theme paradigms could favor the hypothesis of analogical extension in the case of the change *-ud-* > *-id-*, as part of a more general trend for a partial merger of the 2nd and 3rd conjugations since the loss of *-ud-* endings reduced the number of contrasting forms. It should be noted, however, that the systematic character of the change, with no residual verb endings in *-ud-*, is somewhat atypical of analogical changes, which tend to apply only to a subset of the potential “candidate” forms. The exceptionless character of the change would, thus, point to a combination of extension and leveling factors, a combination that may be related to more regular analogical changes (Hock 1991: 179).

Another distinguishing feature between Old and Middle Portuguese, also concerning verb inflexion, is the realization of the 2nd person plural suffixes. The almost uniform realizations *-de* / *-des* (imperative/other inflected forms), with only the exception of the simple perfect (*-stes*), predominant in the former stage, contrast with the variation observable in Middle Portuguese, which demonstrates the loss of intervocalic *-d-*.

As can be observed in Table 1.2, the increased allomorphy in person/number marking resulted not only from divergent realizations that emerged as a consequence of *-d-* loss in the various conjugations (diphthong or vowel through *-i* + *i*-reduction) but also from the preservation of *-d-* in non-intervocalic contexts and after a nasal. Moreover *-d-* was preserved intervocalically in a set of “exceptional” forms. This was explained by scholars on the basis of different factors, phonological (Williams 1975: §155, 4) as well as morphological (Piel 1989: 218). In Middle and Classic Portuguese, there are attestations of variation in both intervocalic (*ides/is*) and non-intervocalic contexts (*fazerdes/fazeres*) that could indicate that *-d-* loss may have been somewhat less restricted initially.

Note that while in contemporary European Portuguese 2nd person plural forms (with or without expressed *vós* “you-plural”) are used only in the northern and in some central dialects with a plural addressee, they occurred regularly, as in other Romance areas, also to

**Table 1.2** Examples of the evolution of second person plural verb suffixes.

| Verbs                 | Old Portuguese (pres.<br>indicative and inflected<br>infinitive) | Results of<br>intervocalic<br>-d- loss | Preserved -d-    |               |
|-----------------------|------------------------------------------------------------------|----------------------------------------|------------------|---------------|
|                       |                                                                  |                                        | after -r-/nasal  | intervocalic  |
| <i>amar</i> "love"    | <i>amades</i><br><i>amardes</i>                                  | <i>amais</i>                           | <i>amardes</i>   |               |
| <i>fazer</i> "do"     | <i>fazedes</i><br><i>fazerdes</i>                                | <i>fazeis</i>                          | <i>fazerdes</i>  |               |
| <i>ter</i> "have"     | <i>tendes</i>                                                    |                                        | <i>tendes</i>    |               |
| <i>crer</i> "believe" | <i>credes</i>                                                    |                                        |                  | <i>credes</i> |
| <i>partir</i> "leave" | <i>partides</i><br><i>partirdes</i>                              | <i>partis</i>                          | <i>partirdes</i> |               |
| <i>ir</i> "go"        | <i>ides</i>                                                      |                                        |                  | <i>ides</i>   |

address a singular interlocutor (a use that persisted until the late eighteenth/ early nineteenth century), thus contrasting with 2nd person singular (*tu* "you-singular"), in distinct pragmatic contexts. The disuse of these forms has been related diachronically to the emergence of honorifics in the late fifteenth/ early sixteenth century (*Vossa Mercê*, *Vossa Alteza*). These were used with 3rd person verb forms (Cintra 1986: 18), expanding later to other contexts. We can only hypothesize that the increased allomorphy in the person/ number verb inflexion that resulted from -d- deletion may also have contributed to the tendency towards the disuse of the 2nd person plural forms, complementing Cintra (1986: 31) who suggests that the disuse of *vós* may have been favored by a tendency to "simplify" the verb inflexion.

### 3. The historical dimension of Brazilian Portuguese

An overview of the historical evolution of Portuguese from the medieval period to the period of Brazilian Portuguese will be based on a periodization already argued for in the literature by other scholars (see Ramos and Venâncio 2006 and references cited there). The dynamics of the societal formation of Brazil, which, from the onset, involved contact between Portuguese speakers and peoples who spoke non-Indoeuropean languages, may be held responsible for the differentiation between European Portuguese (henceforth EP) and Brazilian Portuguese (henceforth BP).

In the remainder of this section we will attempt to put the creation of the linguistic diversity in Brazil in the context of the country's complex social history. This will also help us to understand the actual differences between BP and EP and the different variational aspects of the two varieties.

#### 3.1. The historical origins: Brazilian Portuguese

The social history of Brazil is profoundly marked by the contact established among at least three historical agents: the Portuguese who arrived on the coast of Brazil as of 1500, the indigenous populations, and the African slaves brought to Brazil until late in the nineteenth century (Mattos e Silva 2004b). We distinguish three different periods: 1) the first period of

colonization during the sixteenth century, when the Portuguese language was transplanted, though only to a limited degree, to the Brazilian territory; 2) the particular language setting in the eighteenth century; and 3) the linguistic diversification that took place in the nineteenth century.

The first period represents the beginning of the colonization of Brazil by Portugal in 1532. In the so-called “colonial period” (Teyssier 1982), settlers established themselves on the Atlantic coast. The earliest texts written in Brazil, such as travel stories, historical or geographical writings, treaties, and letters, were mainly descriptive and produced by the newly arrived Portuguese, who were astonished by the nature and the native people of the new world. European settlers represented a minority in Brazil until the eighteenth century. According to Mussa (1991), the Portuguese and their descendants comprised less than 30 percent of the population in the seventeenth century. The indigenous presence was very large; at the beginning of colonization the region’s population was almost five million. Some authors (Lobo 2003, Mattos e Silva 2004a) characterize first three centuries of the social history of Brazil as one of “generalized multilingualism.” Such a designation accounts for the fact that Portuguese was in contact with a large number of indigenous languages, mainly belonging to the Macro-Jê linguistic stock and, on a smaller scale, to the Tupi-Guarani language family. Communication between the Portuguese and the natives was through the “*língua geral*” (widely spoken language) (Rodrigues 1996). The term “*língua geral*” is used to designate two different languages: 1) “*língua geral paulista*,” a Tupi-based language that was spoken in São Paulo and used by the Portuguese explorers, who penetrated the Brazilian hinterlands in search of especially gold and silver or Indians for enslavement, and taken to the central part and south of the country (Goiás, Mato Grosso and the north of Paraná); 2) “*língua geral amazônica*,” a Tupinambá-based language, spoken in the north-northeast (Pará and Maranhão) and taken to Amazonia. Before the arrival of the Portuguese, the Tupi and Tupinambá languages were spoken over a large geographical area. For this reason, they were used for the evangelization of the indigenous populations by the Jesuits during the sixteenth and seventeenth century and became the contact languages used between Portuguese and Indians of different tribes. The linguistic contact between the “*língua geral*” and EP in colonial Brazil constitutes the social-historical basis of BP.

In addition to the contact with indigenous languages, the “generalized multilingualism” of the first period of Brazil was affected by the presence of African languages. Almost five million slaves with distinct languages from different regions of Africa were brought to Brazil in this first period. The Africans and their descendants were the main diffusing agents of BP in its popular variety (Mattos e Silva 2004b). In demographic terms, the African slaves outnumbered the Portuguese settlers by around 60 percent in the seventeenth century (Mussa 1991). The African slaves needed to adopt the settlers’ language in their daily life, learning it as their 2nd language. Their influence was decisive in the linguistic restructuring of non-Standard Portuguese and its diffusion throughout the Brazilian territory (Lobo 2013: 23).

The transition between the seventeenth and eighteenth century was characterized by an increased influence of the African and Portuguese populations on the indigenous peoples and the restriction of the use of the “*língua geral*” as a means of communication with the Indians. Other factors that contributed to the growing importance of Portuguese as the sole language of communication were the accelerated growth of the Portuguese population in Brazil in the eighteenth century (reaching 800,000) and Pombal’s linguistic policy, which imposed the use of Portuguese in official documents. Up to this period, the texts written in Brazil did not present any striking differences as compared with EP. For this reason, it is referred to as “Common Portuguese” as it was spoken in Lisbon.

The nineteenth century effectively gave birth to a new period with important social changes in the first and second halves of the century. It was the period of the “re-lusitanization” of Brazil (Teyssier 1982), which started when the royal family arrived in Rio de Janeiro in 1808 and, later, when many Portuguese immigrants arrived in Brazil in the 1950s. In the

twentieth century, Brazil gradually ceased to be a rural colony due to the urbanization of the coastal cities, the migration of the rural elites to Rio de Janeiro, capital of the kingdom, and the demographic and economic rise of the urban working class. During that period, European and Asian immigrants started to do the work previously done by slaves and favored the spreading of Portuguese, especially in its popular variety. Indeed, the acquisition of Portuguese as a second language by the newcomers took place in the rural areas with foremen, laborers, and descendants of slaves. Since immigrants from different regions of the world settled in different areas of Brazil, this immigration flow laid the basis for the regional linguistic diversity of BP (Lucchesi 1998).

To sum up, the social dynamics that characterized the colonial period in Brazil set up a “crossbreed Portuguese” that took on features foreign to EP due to the confluence of different languages: EP, indigenous and African languages. In the next section, we intend to provide evidence showing that the regional and social heterogeneity of BP is a result of the Brazilian social-historical formation.

### ***3.2. Linguistic settings in Brazilian Portuguese***

The complexity of the historical formation of BP is evidenced by: (1) the contact established between the Portuguese colonizers, natives, and Africans; (2) the Jesuit evangelization of Indians, allowing intercommunication on the coast with the explorers from São Paulo; (3) the diffusion of popular Portuguese carried out by people of African descent (internal traffic); (4) European and Asian immigration; and 5) the internal migration from rural areas to the cities in the twentieth century. These factors produced dialectal heterogeneity at the horizontal and vertical levels, most of all in terms of phonetic–phonological phenomena.

Although there are polarized norms (standard and vernacular) in current BP, some phonetic–phonological phenomena may evidence linguistic differences and similarities between BP and EP. On one side, BP preserves features from the colonial times, while, on the other, it presents specific innovations.

## **4. European Portuguese and Brazilian Portuguese: main contrasting features and changes**

### ***4.1. Some phonetic–phonological distinctions***

As in its origins, when Portuguese distanced itself from Castilian through its vocalism, the modern EP vowel system presents innovations that do not appear in BP and vice-versa. According to Mateus and d’Andrade (2000: 2), “The most obvious differences between these two varieties are located in the unstressed vowel system—the vowels are more audible in BP than in EP.” In final position, for example, the unstressed mid vowels /e/ and /o/ became high vowels around 1800 in EP: [e] > [i] and [o] > [u]. Word-final unstressed [i] almost became silent: bat[e] > bat[i] > bat[i]/bat “beat” (Teyssier 1982). Except after palatalized coronal stops, word-final /i/ does not generally delete in BP, which is considered by some a conservative feature of this variant. Equally, unstressed /a/ in word-final position does not seem as reduced in BP as it is in EP.

While EP has a strong tendency to reduce and delete vowels in unstressed positions, in BP unstressed vowels are generally maintained. However, different from EP, BP shows a tendency towards consonantal weakening and loss in the syllable coda, where the consonants /l, r, s/ suffer vocalization or even total deletion. The behavior of the consonants in coda position distinguishes the two varieties and evidences a phonetic–phonological aspect with morpho-syntactic consequences: the change of a CVC type syllable structure to a CV type.

The deletion of /r/ in word-final position, for example, which occurs only in BP, has a greater incidence in verbs, where this consonant is the infinitive marker: *cantar* > *cantá* “to sing.” In the other lexical categories, the deletion rate of /r/ remains below 40 percent: *colar* > *colá(r)* “necklace” (Callou, Leite, and Moraes 2002). The deletion of /s/, when it occurs word-finally, interferes with nominal agreement and the expression of plurality. In BP, nominal agreement corresponds to a variable rule with a tendency to mark agreement only in the first constituent of a noun phrase, whereas EP expresses the plural affix *-s* on every constituent of the NP that can receive inflection: *os meninos pequenos* (EP) vs. *os meninoØ pequenoØ* (BP) “the little boys” (Scherre 1988; Brandão 2013). This difference in behavior can be associated to the historical formation of BP, since many African languages that came to Brazil, as well as the Tupi languages, did not present a closed syllable (CVC) such as occurs in EP which tends to preserve the consonants in the syllable coda. Although a correlation between external and internal history can be overly reductive, it is undeniable that the complex social-historical formation of BP has set up conditions such that (1) heterogeneity emerged, and (2) the implementation of certain changes propagated more quickly in the new territory.

## 4.2. Some morphosyntactic aspects

The differentiation between BP and EP is less clear-cut in the morphosyntax than it is in the phonology. Existing written documents provide some evidence of different directions of change that have operated probably since the eighteenth century in both areas.

The use of the gerund (*escrevendo* “writing”), nowadays generalized in Brazil, was a recurring strategy to indicate the progressive aspect until the sixteenth or seventeenth century. In EP, the innovative strategy (called the *gerundial infinitive*) formed by combining the preposition *a* “to” with the infinitive (*a escrever* “to write”) started to be used from the eighteenth century on (Barbosa 2008) in variation with the gerund (*escrevendo* “writing”), although the first occurrences of this construction can already be found as early as the fifteenth century. In the eighteenth century, no clear differences in the use of this construction can be observed in BP and in EP. The use of the *gerundial infinitive* starts to increase in Portugal at the beginning of the twentieth century, combined with the auxiliary verbs *estar* “be,” *andar* “walk,” *ficar* “stay” and *continuar* “continue” (Mothé 2007). Studies show different frequency levels depending on the type of text.

The behavior of *haver* and *ter* “have” in existential constructions also shows differences between the two varieties of Portuguese. In terms of its Latin origin, the form *habere* had as its first meaning “to have in one’s possession,” “to store,” while the verb *tenere* meant “to hold in one’s hand,” “to obtain.” Over time, *habere* underwent an extension of its semantic content and assumed the figurative meaning of “to have in one’s hand.” In the first Portuguese documents of the thirteenth century, the verb *haver* used to be more productive for any semantic value of possession as a full verb (Mattos e Silva 1997). The verb *haver* was also used in variation with *ter* for the more concrete meaning of material possession: *haver/ter pan, casa* “to have bread, a house.” In contexts of immaterial possession, when the value of possession was more abstract, the use of the verb *ter* was less productive. As an example, if the possessed object presented “immaterial acquirable properties” the use of *ter* was rare in the thirteenth century (*haver/ter fé* “to have faith”), but, in contexts of “the possessor’s inherent properties,” the verb *haver* was categorical in this period: *haver enfermidade, cegueira* “to have an illness, blindness” (Mattos e Silva 2002). There is some evidence that, until the fourteenth century, the two verbs could convey distinct values with *haver* marking an inherent relation or strict possession (in cases of material possession) and *ter* a non-inherent relation, or, in cases of material possession, a temporary possession.

Gradually, until the middle of the sixteenth century, the verb *ter* took the place of *haver* for all the possession types mentioned, while *haver* had specialized its meaning in existential

constructions and eliminated the etymological form *ser* (verb “be” meaning existence). The generalization of *haver* as an impersonal verb was also favored by the fact that this verb assumed inanimate subjects consisting of place names from the time of Latin: ARCA NOE HABUIT HOMINES “Noah’s ark had men” (Bourciez 1956: 252). These constructions made an existential interpretation possible due to the reanalysis of the subject as a locative complement introduced by the preposition “in”: IN ARCA NOE HABUIT HOMINES “There were men in Noah’s ark.”

In BP, the verb *ter* assumed an existential meaning replacing the verb *haver* in this context probably from the nineteenth century on. However, this replacement only occurs in written texts in the first half of the twentieth century, as shown by Marins (2013). According to Callou and Avelar (2007), the emergence of the verb *ter* as an existential construction may be associated with other linguistic changes in BP. The weakening of the verbal agreement and the inability to interpret a sentence in which the verb *ter* occurs with a null subject as a possessive construction favored its reanalysis as an existential construction. For instance, a sentence such as *Tem várias maçãs na geladeira* “Have several apples in the fridge” is interpreted as “There are several apples in the fridge” in BP, while its interpretation is “He/She has several apples in the fridge” in EP.

The postverbal (V-CI) or preverbal (CI-V) pronominal clitic placement is another aspect recurrently mentioned to distinguish between EP and BP. Martins (2005) argues that medieval Portuguese presents variation between enclisis and proclisis in contexts that are ungrammatical in modern Portuguese. This variable behavior—preverbal in (1) or postverbal in (2)—occurred in affirmative main clauses without triggers of proclisis (negative items, *wh*-phrases, quantifiers, etc.), that is, in the “unmarked main clauses”:

- (1) Sobrinho, eu *vos* rogo que fiques aqui  
Nephew, I *you-pl-dat* beg that stay here  
“Nephew, I beg you to stay here.”
- (2) Rogo-*vos* que nom vaades em esta demanda  
I-beg-*you-pl-dat* that not go in this demand  
“I beg you not to make this demand.”  
(Fifteenth century *A Demanda do Santo Graal*)

Written texts in the medieval and early Renaissance period show a progressive increase in the use of clitic–verb order. In the thirteenth century, enclisis was predominant in the unmarked main clause; in the sixteenth century, however, proclisis became more frequent than enclisis. Although the verb–clitic order was infrequent during the sixteenth century, it was grammatically possible, mainly in neutral sentences. Martins (2005) argues that proclisis was emphatic and semantically motivated in EP in the period extending from the thirteenth to the sixteenth century, while enclisis was interpreted as neutral and non-emphatic. From the seventeenth century on, the verb–clitic order replaced the sequence clitic–verb, just as in modern EP.

Thus, it is possible to say that the Portuguese settlers had brought this variation in the preverbal and postverbal placement of clitic elements in main clauses to Brazil. In BP, however, the clitic–verb order became dominant in contexts where enclisis is obligatory in EP. For instance, we can see proclisis in BP in the first absolute position as in (3), the co-occurrence of clitics with past participle in (4) and gerunds in (5) in verbal periphrastic constructions:

- (3) Melembro muito de você com muitas saudades  
*Me-dat-remember.pres.1sg* very much of you with very miss-noun  
“I remember you very much and miss you so much.”  
(Missive written by Barbara Ottoni, nineteenth century. In Lopes, 2005: 215)

- (4) ... verificar que tens te adiantado  
... to verify that you-have-NOM you-acc attempted  
“... to verify that you have attempted yourself”  
(Missive written by Barbara Ottoni, nineteenth century. In Lopes 2005: 127)
- (5) estou sempre melembrando de que você sempre queria  
I-am always me-refl-remembering of that you always wanted.2sg  
me ajudar  
me-acc-help.inf  
“I always remember that you always wanted to help me.”  
(Missive written by Barbara Ottoni, nineteenth century. In Lopes 2005: 215)

The facts, as described above, reconstructed on the basis of the remaining BP documentation, show that, in BP, clitic placement is sensitive to external factors, such as the type of text, adoption of standard EP, social prestige, etc. In addition, Brazilian texts of the nineteenth century contain some examples of clitic–verb orders very common nowadays in spoken BP.

## 5. Conclusions

The preceding sections provide a diachronic overview of the Portuguese language. We have briefly reviewed some of the more salient tendencies, both conservative and innovative, which determine the Portuguese differentiation within the Ibero-Romance context. This perspective has also guided the description of the phonetic-phonological and morpho-syntactic phenomena which contributed to distinguish BP from EP. As shown, the differences between these two varieties are more obvious and outstanding in the phonology than in the morphosyntax. The most striking discrepancy is probably to be found in the vocalism, especially the unstressed vowel system, which is subject to different neutralization strategies in EP and BP. From the seventeenth century on, BP texts do not show significant morphosyntactic differences as compared with EP, which is due to the fact that a very small literate group was responsible for writing the extant texts. Nevertheless, it could be shown that, regarding the phenomena selected here, as from the eighteenth and the nineteenth century, EP and BP took separate paths. Detailed analyses of the synchronic differences between BP and EP phonology, morphology, and syntax can be found in other chapters of this volume.

## REFERENCES

---

- Baldinger, K. (1972). *La formación de los dominios lingüísticos en la Península Ibérica*, 2nd edn. Madrid: Gredos.
- Barbosa, A. G. (2008). Fontes escritas e história da língua portuguesa no Brasil: as cartas de comércio no século XVIII. In I. Lima and L. do Carmo (eds.), *História social da língua nacional*. Rio de Janeiro: Edições Casa de Rui Barbosa, pp. 181–214.
- Bourciez, E. (1956). *Elements de Linguistique Romane*, 4th edn. Paris: Librairie C. Klincksieck.
- Brandão, S. F. (2013). Patterns of agreement within the noun phrase. *Journal of Portuguese Linguistics*, 12 (2), pp. 51–100.
- Callou, D. and Avelar, J. (2007). Sobre a emergência do verbo possessivo em contextos existenciais no português brasileiro. In A. Castilho et al. (eds.), *Descrição, História e Aquisição do português brasileiro*. São Paulo: Pontes/FAPESP, pp. 375–402.
- Callou, D., Leite, Y., and Moraes, J. (2002). Processo(s) de enfraquecimento consonantal

- no português do Brasil. In M. B. Abaurre and A. Rodrigues (eds.), *Gramática do português falado VIII: novos estudos descritivos*. Campinas: UNICAMP/FAPESP, pp. 537–555.
- Câmara, J. M. (1975). *História e estrutura da língua portuguesa*, 2nd edn. Rio de Janeiro: Padrão.
- Castro, I. (2006). *Introdução à história do português*, 2nd edn. Lisboa: Colibri.
- Cintra, L. F. L. (1983). *Estudos de Dialectologia Portuguesa*. Lisboa: Sá da Costa.
- Cintra, L. F. L. (1986). *Sobre “formas de tratamento” na língua portuguesa*, 2nd edn. Lisboa: Livros Horizonte.
- Hock, H. H. (1991). *Principles of Historical Linguistics*, 2nd edn. Berlin: Mouton de Gruyter.
- Lobo, T. (2003). A questão da periodização da história linguística do Brasil. In I. Castro and I. Duarte (eds.), *Razões e emoção. Miscelânea de estudos em homenagem a Maria Helena Mira Mateus*, Vol. 2. Lisboa: Imprensa Nacional-Casa da Moeda, pp. 395–411.
- Lobo, T. (2013). Rosa Virgínia Mattos e Silva e a sócio-história do português brasileiro. <http://www.usc.es/revistas/index.php/elg/article/view/2314/2861>, accessed October 10, 2015. DOI: <http://dx.doi.org/10.15304/elg.7.2314>
- Lopes, C. R. S. (2005). *A Norma brasileira em construção: fatos linguísticos em cartas pessoais do século XIX*. Rio de Janeiro: FAPERJ/PPGLEV.
- Lucchesi, D. (1998). A constituição histórica do português brasileiro como um processo bipolarizado: tendências atuais de mudança nas normas culta e popular. In S. Große and K. Zimmermann (eds.), *“Substandard” e mudança no português do Brasil*. Frankfurt am Main: Teo Ferrer de Mesquita (TFM), pp. 73–100.
- Maia, C. A. (1986). *História do galego-português: Estado linguístico da Galiza e do Noroeste de Portugal do século XIII ao século XVI*. Coimbra: INIC.
- Maiden, M., Smith, J. C., and Ledgeway, A. (eds.). (2011). *The Cambridge History of the Romance Languages*, Vol. I *Structures*. Cambridge: Cambridge University Press.
- Marins, J. E. (2013). As repercussões da remarcação do Parâmetro do Sujeito Nulo: um estudo diacrônico das sentenças existenciais com *ter* e *haver* no PB e no PE. Rio de Janeiro: Federal University of Rio de Janeiro (PhD Thesis).
- Martins, A. M. (1999). Ainda “os mais antigos textos escritos em português”: Documentos de 1175 a 1252. In I. H. Faria (ed.), *Lindley Cintra: Homenagem ao Homem, ao Mestre e ao Cidadão*. Lisboa: Cosmos / FLL, pp. 491–534.
- Martins, A. M. (2001). On the origin of the Portuguese inflected infinitive: A new perspective on an enduring debate. In L. J. Brinton (ed.), *Historical Linguistics 1999: Selected Papers from the 14th Conference on Historical Linguistics*. Amsterdam: John Benjamins, pp. 207–222.
- Martins, A. M. (2005). Clitic placement, VP-ellipsis and scrambling in romance. In M. Batllori et al. (eds.) *Grammaticalization and Parametric Change*. Oxford: Oxford University Press, pp. 175–193.
- Mateus, M. H. M. and d’Andrade, E. (2000). *The Phonology of Portuguese*. Oxford: Oxford University Press.
- Mattos e Silva, R. V. (1997). Observações sobre a variação no uso dos verbos *ser*, *estar*, *haver*, *ter* no galego-português ducentista. *Estudos lingüísticos e literários*, 19, pp. 253–285.
- Mattos e Silva, R. V. (2002). Vitórias de *ter* sobre *haver* nos meados do século XVI: usos e teorias em João de Barros. In R. V. Mattos e Silva and A. Venâncio Filho (eds.), *O Português Quinhentista—Estudos Lingüísticos*. Salvador: EDUFBA/UEFS, pp. 121–142.
- Mattos e Silva, R. V. (2004a). *Ensaios para uma Sócio-História do Português Brasileiro*. São Paulo: Parábola Editorial.
- Mattos e Silva, R. V. (2004b). O português brasileiro: sua formação na complexidade multilinguística do Brasil colonial e pós-colonial. In S. B. B. Costa, A. Machado Filho, and A. V. Lopes (eds.), *Do português arcaico ao português brasileiro*. Salvador: EDUFBA.
- Mothé, N. G. M. (2007). *Variação e mudança aquém e além mar: gerúndio versus infinitivo gerundivo no português dos séculos XIX e XX*. Rio de Janeiro: UFRJ (MA thesis).
- Mussa, A. B. N. (1991). *O papel das línguas africanas na história do português do Brasil*. Rio de Janeiro: UFRJ (MA thesis).
- Piel, J. M. (1989). *Estudos de Linguística Histórica Galego-Portuguesa*. Lisboa: IN-CM, pp. 213–244.
- Ramos, J. M. and Venâncio, R. P. (2006). Por uma cronologia do português escrito no Brasil. In T. Lobo et al. (eds.), *Para a História do Português Brasileiro*. Salvador: EDUFBA, pp. 575–595.
- Rodrigues, A. (1996). As línguas gerais sul-americanas. *Papia*, 4 (2), pp. 6–18.
- Sampson, R. (1999). Ibero-Romance II: Galician-Portuguese. In *Nasal vowel evolution in*

- Romance*. Oxford: Oxford University Press, pp. 175–218.
- Scherre, M. P. (1988). *Reanálise da concordância de número em português*. Rio de Janeiro: UFRJ (PhD thesis).
- Teyssier, P. (1982). *História da Língua Portuguesa*. Lisboa: Sá da Costa.
- Vincent, N. (1988). Latin. In M. Harris and N. Vincent (eds.), *The Romance Languages*. London: Routledge, pp. 26–78.
- Williams, E. B. (1975). *Do latim ao português. Fonologia e morfologia históricas da língua portuguesa*, 3rd edn. Rio de Janeiro: Tempo Brasileiro.