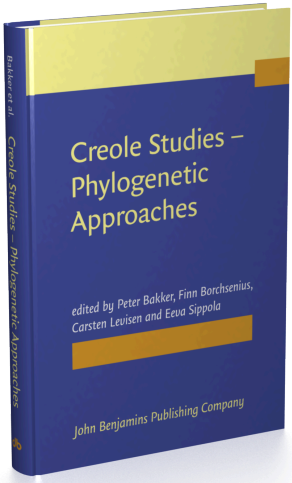
CHAPTER 11

Similarities and differences among Iberian creoles

Eeva Sippola | University of Bremen

https://doi.org/10.1075/z.211.11sip



Pages 241–268 of
Creole Studies – Phylogenetic Approaches
Edited by Peter Bakker, Finn Borchsenius, Carsten Levisen and Eeva M. Sippola
2017. x, 414 pp.

CHAPTER 11

Similarities and differences among Iberian creoles

Eeva Sippola
University of Bremen

This study provides an account of the classifications of Iberian creoles based on a phylogenetic network analysis of typological and lexical data. It maps the shared linguistic traits of these creoles and compares the differences between them. The results of the typological analysis support previous classifications, showing a clear division between Atlantic and Asian creoles, and the main subgroups are divided geographically. However, connections between the subgroups are less clear due to the degree of detail of the typological feature values. The similarities consist of typologically unmarked features that are shared with the lexifiers. The differences are a result of areal substrate and adstrate influence.

Keywords: Iberian creoles, areality, groupings, similarities, differences

11.1 Introduction

Creoles with a lexicon drawn from Spanish and Portuguese represent the oldest, most diverse, and most geographically dispersed creoles. This chapter provides an account of the classifications of this subgroup of contact varieties based on a phylogenetic network analysis of typological and lexical data. It aims to map the shared linguistic traits of the Iberian creoles and compare the differences between them.

Previous comparative studies of Iberian creoles (e.g. Cardoso et al. 2012; Clements 2009; Hagemeijer 2011; Jacobs 2012) have identified clusters based on geography and on structural and lexical features. However, most of these studies have concentrated on individual languages, areal subclusters, and features characteristic of a subgroup of creoles. To date, few studies have taken a wide comparative focus on these languages (for exceptions, see Hancock 1975; Ferraz 1987). The diversity of the varieties and the complex linguistic situations in which the creoles formed have made it challenging to offer a comprehensive account of shared features, although differences have been pointed out. To this we can add problems concerning the scarcity of the data available for many creole varieties.

DOI 10.1075/z.211.11sip

Recently published databases and newly available computational methods offer the possibility of a wide comparative approach such as the one taken in this study.

This study analyses both grammatical and lexical traits of 19 Iberian creoles. The data come mainly from the *Atlas of Pidgin and Creole Language Structures Online* database (Michaelis et al. 2013). The analysis is based on statistical modelling and computational tools for quantitative typology that are suitable for language contact situations, such as the phylogenetic programs SplitsTree and Mesquite (Daval-Markussen & Bakker 2011, 2012; Huson & Bryant 2006; Maddison & Maddison 2014). More generally, the phylogenetic classification of English-based Atlantic creoles by Daval-Markussen & Bakker (2011) has served as an inspiration for this paper. The focus of this study, however, is on the shared features and differences of languages from a wider geographical area, covering both Atlantic and Asian creoles.

Section 11.2 offers comments on the methods used in this chapter. Section 11.3 gives an overview of previous classifications of Iberian creoles and discusses some of the shared and differing features that have been identified among them. Sections 11.4 and 11.5 present the language sample and the typological features used in the analysis. Sections 11.6–11.8 cover the results, including the areal clusters of the creoles and the shared features and differences among them, followed by a discussion in Section 11.9. The conclusion follows in Section 11.10.

11.2 Methods

Both distance-based and character-based methods of phylogenetic analysis are used in this study. The software SplitsTree4 (Huson & Bryant 2006) is used to provide an overview of the classification of the Iberian creoles, their groupings, and the number of conflicting signals. The program creates phylogenetic networks using a distance-based method called Neighbor-Net (Bryant & Moulton 2004). First, the primary typological and lexical feature values are converted into a numerical matrix. Then the software is used to calculate the difference between each language in the sample by means of combining all the differences and similarities between the languages into a single distance measure. The resulting unrooted network (split graphs) is thus only based on language similarities and differences, and does not embody phylogenetic models as such (Wichmann et al. 2011: 207; Verkerk 2014: 12–13). In this case, the internal nodes of the graph represent the conflict between different splits in the data analysis. The conflict of the signal is produced in the contradictory groupings of taxa. The internal nodes do not indicate ancestors, and they do not point to specific features that are responsible for the conflicting signals (List et al. 2013: 145).

In addition to an impressionistic interpretation of the networks, we are interested in measuring the reticulation in them in a principled way. Consequently, auxiliary methods tested in this study include the calculation of delta scores in order to detect the involvement of each variety in the conflicting signal and to assess if this method can be applied to our empirical sample. According to Gray et al. (2010: 3925), the delta score for each taxon measures to what extent each language connects with the conflicting signal. It ranges from 0 to 1, and equals zero if the language is not involved in any conflicting signal (Wichmann et al. 2011: 210 give a more specific explanation of the calculation of the score).[1]

As a consequence of using distance-based methods, we cannot identify the characters or features that cause conflict in the networks (List et al. 2013: 145). Moreover, we cannot be sure that the clustering patterns are due to inherited features or common ancestry and not parallel development or borrowing. Therefore, a supplementary analysis was carried out using maximum parsimony implemented in PAUP* *Phylogenetic Analysis Using Parsimony 4.10.b* (Swofford 2002). Maximum parsimony analysis constructs the tree that requires the fewest changes in the values of each feature to evolve from an ancestor to all the observed descendants or language varieties (Dunn 2009: 1667). The strict consensus tree (here interpreted as a classification of synchronic features) was used to calculate ancestral states in order to identify those features that appear to be unique to certain clusters and those that have likely evolved independently or that have been subject to lateral transfer/borrowing. This mapping of the features was carried out with the help of the program *Mesquite Tools for Phylogenetic Trees* (Maddison & Maddison 2014).

11.3 Previous classifications

Iberian creoles have been studied since the early days of creolistics. Earlier research has focused mostly on individual languages, connections between a restricted number of languages, and areal subclusters. Here we will present the relevant findings for our comparative purposes. For historical and current information on Iberian creoles (speaker numbers, period of formation, and other sociohistorical information) we refer the reader to the descriptions in Volume II of APiCS (Michaelis et al. 2013).

---

1. A Q-residual score is a similar measure for calculating departures from a strict tree. However, it is dependent on scaling of the distances meaning that the lengths of terminal branches play a role. Gray et al. (2010: 3925) consider it to provide a more accurate value closer to the residual in standard statistics, but Wichmann et al. (2011: 212) note that the distinctiveness of one or more taxa in relation to the others should not affect the measure.

11.3.1 Early studies

Early creolists such as Adolpho Coelho and Hugo Schuchardt published work on Portuguese-based and/or Spanish-based creoles as far back as the decade between 1880 and 1890. Coelho can be characterized as an early universalist, pointing towards the mechanisms of imperfect acquisition in creole formation (Coelho 1881: 67, 69). Coelho supported his theory by pointing out similarities among Portuguese-lexifier creoles, such as the progressive preverbal marker *ta* and a number of shared lexical items, such as *miste* 'need' and *papia* 'speak' (Coelho 1881: 70; Holm 1988: 28). The father of modern creolistics, Schuchardt, published widely on Iberian creoles. In *Kreolische Studien* (Schuchardt 1882a, 1882b, 1883a, 1883b, 1884, 1888d, 1890) and other writings on Portuguese contact varieties (Schuchardt 1888a, 1888b, 1888c, 1889a, 1889b, 1889c), Schuchardt studied creoles from places ranging from São Tomé to the Philippines, grouping them based on their substrate languages and geographical areas. However, he tried to account for particular creole features individually, as a result of how they arose in each contact situation (Holm 1988: 30). The third early creolist of special relevance for Iberian creoles is Rodolfo Sebastião Dalgado, who collected Indo-Portuguese texts around the turn of the 20th century. He was the first to propose a theory of partial reciprocal transfusion for Portuguese-based creoles, where different varieties exerted mutual influence on each other. His hypothesis maintains that the Asian creoles were able to develop and maintain significant similarities to each other because of the economic and political links between the territories under Portuguese rule, which facilitated cultural interchange and population movements between them (Cardoso 2010).

The rapid growth of creole research that began in the 1960s included several studies on Iberian varieties and their connections to other creole varieties. A hypothesized early Portuguese pidgin played a significant role in the monogenetic theories of creole formation (e.g. Taylor 1971; Thompson 1961; Whinnom 1956, 1965). According to these views, the common structures shared by creoles were due to their origin in a Portuguese proto-pidgin that was used in West Africa and in Portugal and later carried to different parts of the world along with colonial expansion. This Portuguese pidgin would have been later relexified during the development of the English, French, Spanish, and Dutch creoles. The linguistic evidence of this variety is scarce, but varieties of this kind can nevertheless be shown to have existed based on observations preserved in literary works of the period (see, for example, Naro 1978; Kihm & Rougé 2013).

Another example of the monogenetic approach can be found in Whinnom's account (1956) of the formation of the Philippine Spanish creoles, also known as Chabacano. He postulated that the Chabacano varieties were based on an earlier Asian Portuguese pidgin, which was later relexified by Spanish and taken to the

Philippines. It is true that certain similarities can be identified between Chabacano and the Portuguese-based creoles of Southeast Asia, especially those in Malacca and Macau. For example, these creoles have similar pronouns, a formally and functionally similar object marker *ku/ko* or *kon/kun(g)*, and a similar verbal marker *kaba* with completive functions. However, these shared syntactic structures are largely universal tendencies in creoles that have two closely related lexifiers, and no irrefutable linguistic evidence exists to account for a historical connection between a previous Portuguese pidgin and the Chabacano varieties spoken today (Clements 2009: 48, 53; Lipski 1988: 37). For example, Chabacano words that have been assigned a Portuguese origin, such as *agora* 'now' and *onde* 'where', are also attested in Old Spanish, and others, such as *endenantes* 'before' and *boneca* 'doll', are also found in contemporary dialects of Spanish (Lipski 1988: 34–35). Items that are more difficult to classify as Portuguese, Spanish, or substrate derived are *na* 'LOC', the concessive *maskin* 'even if' (but see Vazquez Veiga & Fernández 2012), *quilaya* 'how', and *cosa* 'what' (Lipski 1988: 35–36).

Similarly, arguing for a monogenetic base in the formation of Papiá Kristang, Hancock (1975) gave an overview of the different historical and grammatical connections between 14 Iberian-based creoles and a list of shared features and differences between them. The following creoles were included in the comparison: Papiá Kristang, Macau creole, Java creole, Sri Lanka Portuguese (Ceylon), Bombay creole, the creole of Annobon, the creole of São Tomé, Angolar, Principense, Guinea-Bissau Kriyol, Cape Verdean, Papiamentu, Palenquero, Ermiteño (a variety of Chabacano), Saramaccan, and the pidgin Lingua Franca. Hancock's overview of the features was based largely on Taylor's (1971: 294–295) list of creole features. He identified shared Asian and African features, such as the formation of a past participle with – *du*, the prepositions *na* and *ku*, and similarities across the TMA markers. He also observed differences in the combinatory possibilities of the TMA markers and in the use of the third person plural pronoun, which can function as a post-nominal pluralizing morpheme in some of the Atlantic varieties (Hancock 1975: 221–222). Hancock (1975: 214–215) stated that "the different characteristics of each Lusoasian, Lusoafrican and Lusocaribbean Creole may be attributed to each becoming depidginized in different linguistic environments, the common core remaining unaffected in each".

11.3.2 Atlantic and Asian creoles

Using early creolists' classifications as a starting point, Ferraz (1987: 348) made a first general classification of the Portuguese-based creoles based on structural and lexical features. He established two distinct groups among the Portuguese creoles of West Africa and four groups in Asia. In West Africa, he identified the Upper Guinea

creoles and the Gulf of Guinea creoles. In Asia, the varieties were divided into two Indo-Portuguese groups, Malayo-Portuguese and Sino-Portuguese. Although Ferraz (1987: 339) reasoned that the Portuguese creoles originated independently, he also pointed out the close historical contact between the creole varieties during their formative stages, especially in Asia. This contact during the formative states led to shared grammatical features and lexical items in the Indo-Portuguese varieties (see Dalgado's theory discussed in 11.1). In Africa, Ferraz (1987: 348) described the similarities of the creoles as having arisen from parallel development, with influence from different substrates in the two or four areas mentioned.[2]

The distinction between the Gulf of Guinea and Upper Guinea creoles is based on the features listed in Tables 11.1 and 11.2 (Ferraz 1987: 342–348).

Table 11.1 Differences between the Gulf of Guinea and Upper Guinea creoles (Ferraz 1987: 342–346).

	Feature	Upper Guinea creoles	Gulf of Guinea creoles
Phonetics and phonology	Depalatalization	no depalatalization	depalatalization before /a, o, u/; palatal consonants before /i̲, ĩ, j/
	Word endings	no constraint against consonants	words typically end in a vowel
	Vowel harmony and sandhi rules	no vowel harmony	vowel harmony and sandhi rules present
Grammatical items and functions	Impersonal pronoun *a*	not present	present
	Word order	demonstratives and possessives precede the noun	demonstratives and possessives follow the noun
	Diminutives	preserved	not preserved
	Particle of obligation 'must'	*meste*	*sɛˈla/θɛˈla/ θɛˈra/ʃya*
	Quantifier *muito*	preceding the referent	following the referent
	Serial verbs in complement clauses	no serial verbs in complement clauses	'come', 'go', and 'speak' used in serial verbs

2. This means that the languages in each subgroup show resemblances because they developed together conditioned by similar social and historical processes, controlled by a centralized administration, but with local differences, for example, in the number of speakers of the substrate languages.

Table 11.2 Lexical differences between the Gulf of Guinea and Upper Guinea creoles (Ferraz 1987: 346–348).

Lexical item	Upper Guinea creoles	Gulf of Guinea creoles
'abdomen'	unreduced forms of Ptg. *barriga*	reduced forms, such as *bɛga* and *bwɛga*
'sky'	no prothetic vowel *sɛu*	prothetic vowel *ɔpɛ, ɔsɛ*
'person, someone'	closer to Ptg. *pessoa*: *psoə/ pusua*, *arge/alge*	*nge*, *ninge* (< Ptg. *ninguém* 'nobody')
'to play, to enjoy oneself'	*brinka* < Ptg. *brincar*	*flɔga* and similar forms < Ptg. *folgar*
'pretty, beautiful'	*bunitu*	*g(l)avi*, *fumozu* < Ptg. *formoso*, *nyuka*
'quickly'	*tprɛsə* < Ptg. *depressa*	reduplicated forms, e.g. *nja-nja*
'son, daughter'	*fiye*, *fiya*, *m(e)mine*, *m(e)mina*, and similar forms	*minu*, *mina*, *na*
'stone'	words from Ptg. *pedra*	*budu*, *ubudu*

Turning to Ibero-Asian creoles, Ferraz again pointed out local Asian developments in grammatical and lexical structures (1987: 349–355). Features present only in Asian Portuguese varieties, and not in Africa, include:

1. Particular possessive constructions, such as *Pedro sua casa* 'Pedro's house' and *eu sua casa* 'my house', which are not found in Africa.
2. Modifiers preceding the noun.
3. *-mente* as a productive suffix in adverbs of manner.
4. *Lo* (Indo-Portuguese) and *logo* (Malayo-Portuguese) as future markers.
5. *Ja* as the perfective marker.
6. Reduplication expressing plurality.
7. *Por/para* before a pronominal object in the Indo-Portuguese creoles, and *ku* in Malayo-Portuguese.
8. Portuguese *ter* 'to have' acquires the meaning 'to be' in the creoles.
9. Constructions derived from Old Portuguese *laia* 'way, manner'.

In addition to these grammatical elements, Ferraz (1987: 354–356) provided lists of lexical items characteristic of some or all of the Portuguese varieties in Asia. Table 11.3 presents archaic and nonstandard Portuguese items from Asian creoles not found in West African creoles, while Table 11.4 lists words of substrate origin.

Table 11.3 Archaic and nonstandard Portuguese lexical items in Asian creoles.

(Old) Portuguese item	Creole item
(a) tarde '(in the) afternoon'	SLP[3] *a'tarde*, PK *a'tadi*
a/(à) noite '(at) night'	PK *a'nuti*, TUGU *a'noti*, MAC *a'note*
frio, *frialdade* 'coldness'	SLP *frui'dade*, BAT *friu'dadi*
adem 'duck'	NOR *'ada*, DAM *ad*, SLP *'ade*, BAT *adi*, PK *'a(r)di*
na mais 'only'	SLP *na'ma(i)s*, MAC *na mas*
secura 'thirst, dryness'	NOR *sekur*, SLP *se'kura/se'kuru*, BAT *si'kura*

Table 11.4 Substrate-derived items in Asian creoles.

Substrate item	Creole item
'clay' < Konkani-Marathi *mati* (Dalgado 1921:42)	DAM/GOA *'mate,* MAC *'mati*
'flower' < Indic *phul* 'to bloom'	NOR/MAHÉ/BAT/BAC *'fula,* SLP *'fu:la*
'frog' < Konkani *manduk* 'edible frog' (Dalgado 1921:27)	SLP, PL, BAT, MAC *man'duku/o*
'gift' < Konkani *sâguvât*	DAM *sa'gwat*, SLP *sago'ate/sago'vate*, PK *sa'gwati*
'jacket' < Malay *baju*	BAT, PK, MAC *'baju*
'sad, sadness' < Malay *saying* 'regret', 'pity', 'sorrow for', 'affectionate pining', 'love'	SLP *sa'yan*, MAC *sa'yang*, PK *sa'yang* 'love'
'walking stick' < Malay *rotan*	NOR *rɔt*, SLP, MAC *'rɔta*
'washer man' < Malayalam *mannattan ~mannatti* (Dalgado 1921:12)	NOR *mai'nat*, SLP, MAC *mai'nato*, BAT *mai'natu*
'watermelon' < Arabic *battikh* (Dalgado 1921:191)	NOR *pa'tɛk,* SLP, MAC *pa'tɛka*

Focusing on shared features, Holm (1988–1989: 266–267) examined the verbal markers of the Portuguese-based creoles and concluded that the differences in the patterning of the verbal markers lend support to the African subgroupings established by Ferraz (1987). However, he also pointed out that the creole varieties are interrelated not only through their European lexicon, but that they also share common syntactic features that do not derive from European Portuguese. Holm (2009) examined five additional features of Portuguese-based creoles in Asia that he proposed could originate from the general Portuguese pidgin developing in Africa in the 15th century. These are the verb 'to go' (*vai/vay*), which is derived from an inflected form in Portuguese, and other functional correspondences among African and Asian varieties, such as the completive *kaba*, the coordinating conjunction *ku*, the preposition *na*, and the negator *nunca*.

3. Abbreviations: BAT = Batavia Creole, DAM = Daman Creole, MAC = Makista, NOR = Northern Indo-Portuguese creoles, PK = Papiá Kristang, SLP = Sri Lanka Portuguese, TUGU = Tugu Creole.

11.3.3 Local developments

Today, we have access to a growing number of studies focusing on Iberian creoles. Taking a general approach, Clements (2009) examined two phenomena across the Iberian creoles. These present differences between the African and Asian varieties and among them: affixal tense-aspect markers and oblique, direct object, and indirect object markers. Clements (2009: 50) argued that the presence of affixal features depends on the degree of typological similarity of the languages present in the contact situation, markedness, and social conditions. Suffixal morphology is more likely in creoles that have formed in a more homogeneous contact situation and that have had a stronger presence of the target language, Portuguese (Clements 2009: 55). The oblique and object markers pattern differently between the African and Asian creoles. In the African varieties, the markers follow the predictions of the causal order hypothesis (Croft 1991: 185), in which antecedent roles (e.g. cause, comitative, instrumental) are coded similarly and differently from subsequent roles (e.g. benefactive, dative/recipient), but due to substrate influence, Asian creoles violate this hypothesis (Clements 2009: 67).

Several other studies have been made focusing either on smaller subgroups of Iberian creoles or on certain features. In the Atlantic area, a large number of shared structural and lexical items of African, West Atlantic origin point to a common origin for the Upper Guinea creoles, although the exact location of this origin is unknown (Baptista, Mello & Suzuki 2007; Jacobs 2012: 8; Quint 2000: 99–117). Jacobs (2009, 2012) has established a historical relationship between the Caribbean Papiamentu and the Upper Guinea creoles. This connection is based on the correspondences in morphology and several grammatical categories, such as pronouns, prepositions, question words, conjunctions, reciprocity and reflexivity, TMA markers, and auxiliary verbs. He also points out differences that separate Papiamentu from the Gulf of Guinea creoles and Palenquero (Jacobs 2012: 261). The African connection of Palenquero has also attracted attention, but the origins of this creole are not clear (Jacobs 2012: 13; Lipski 2008; Schwegler & Green 2007). In the Gulf of Guinea, a common origin is generally assumed for the varieties, based on lexical and structural comparisons (e.g. Ferraz 1979; Hagemeijer 2011; Parkvall 2000: 133).

The Asian creoles have received less attention from a comparative perspective. The interrelationships between the Indo-Portuguese varieties and the other Portuguese creoles in Asia have been noted on the basis of early studies (see, for example, Baxter 1996; Clements 2009; and the contributions in Cardoso et al. 2012). Similarly, historical connections on the Malacca-Macau and Malacca-Java axes have been established. Features that are shared along the Malacca-Macau axis include SVO order, accusative-dative markers derived from Portuguese *com*, negative perfective markers of the type *nenang inda* (PK) and *inda nunka* (MAC), vowel + [ŋ] as a reflex of Portuguese nasal vowels, and Malay lexical items (Batalha

1988: 318–319; Baxter 1996: 303–304). In the Malacca-Java area, shared features are verb serialization to express directional motion and give-serialization, and the 1SG.POSS *minya* (for further information, see Baxter 1996: 304). The idea of a possible relationship or shared origin between Chabacano and the Portuguese creoles of Asia surfaces occasionally (for example, Steinkrüger 2006), although a local, independent development of the Chabacano varieties is more widely accepted today (Fernández 2004, 2007, 2012; Lipski 1988, 1992).

To conclude, we notice that previous research has thus established two focal points for Atlantic creoles that have stimulated discussion about the interrelationships between areal subclusters: West Atlantic in the Upper Guinea area and the Bantu and Kwa substrate area in the Gulf of Guinea. In Asia, the emerging picture is on the one hand of substrate influence, and on the other, of partial diffusion between the varieties. In the following sections, we will examine whether computational methods and typological data show support for the groupings and hypotheses presented in the previous literature, and if so, to what extent.

11.4 Language sample

The analysis focuses on the 19 Iberian creole varieties included in the APiCS database. The varieties are (ordered geographically) Palenquero, Papiamentu, Cape Verdean Creole of Brava, Santiago Cape Verdean Creole, São Vicente Cape Verdean Creole, Casamancese Creole, Guinea-Bissau Kriyol, Angolar, Fa d'Ambô, Principense, and Santome in the Atlantic area (see Map 11.1), and Diu Indo-Portuguese, Korlai, Sri Lanka Portuguese, Papiá Kristang, Batavia Creole, Cavite Chabacano, Ternate Chabacano, and Zamboanga Chabacano in Asia (see Map 11.2).

Map 11.1 Atlantic creoles in the sample.

Map 11.2 Asian creole languages in the sample.

11.5 Feature data

The main feature sample for this study is based on the 130 APiCS features. This is the most comprehensive database that includes Iberian creoles. The 130 features are mostly structural and drawn from different areas of grammar: phonology, morphology, syntax, and lexicon. Although the selection is one of convenience, it is partly based on the features included in the *World Atlas of Language Structures* and on features that are typical of contact languages. The features are listed on the APiCS online database, and the division between feature categories employed in APiCS is presented in Table 11.5. In the database, the lexicon and phonology are underrepresented compared to morphology and syntax.

In APiCS, a feature has multiple values (between two and nine, depending on the feature). These values were manually coded into a matrix, using numerical coding between 1–6, meaning that the maximum number of values that could be present for a feature was six (features 94 *Instrument relative clauses* and 104 *Focusing of the noun phrase*).[4] Additional information on the same features was added for Spanish and Portuguese in order to include the lexifiers in the comparison.

4. The difference between the original maximum, nine, and six is due to the fact that not all the APiCS languages are included in this study.

Table 11.5 Percentages of feature categories in APiCS.

Category	Features	
Nominal categories	24	18.5%
Clausal syntax	19	14.6%
Argument marking	16	12.3%
Verbal categories	14	10.8%
Phonology	13	10.0%
Word order	12	9.2%
Lexicon	11	8.5%
Complex sentences	8	6.2%
Negation, questions, focusing	7	5.4%
Nominal syntax	6	4.6%
Total	130	

APiCS offers a solid typological feature set designed especially with creoles in mind. The datasets have been prepared by experts in each language, and the feature values are often based on extensive fieldwork. The atlas offers a wide comparative picture, without the limitations of focusing only on specific areal clusters or lexifier features. Among the Iberian creoles, the sample includes current varieties and one historical variety, namely Batavia Creole. The data on extinct Batavia Creole are based on late 19th and early 20th century texts. The information on the features and languages included in the atlas is rather comprehensive. Information from the APiCS database is missing in only 2.3% of the traits on average. The great majority of the datasets have been completed in their entirety. Angolar, Sri Lanka Portuguese, Diu, and Fa d'Ambô are missing more information than the other languages, most probably due to the precarious situation of these varieties, but even the extinct Batavia Creole has a completion rate of 76.9%. However, as is to be expected, even this database presents some lacunas. There is a lack of information on the highly endangered Indo-Portuguese creoles, such as the Malabar creoles, which have not been described or documented comprehensively so far (but see Cardoso 2013), and Macau Creole Portuguese (Makista). In addition, no Afro-Portuguese or Afro-Hispanic varieties are included (see Perez et al. this volume, Chapter 12, for information on Afro-Hispanic varieties). These are clear limitations, but nevertheless, our study is the largest comparative picture offered so far.

Typological features have been avoided in comparative linguistics with a diachronic approach because they have few possible values and are prone to chance resemblance, their values are linked by implicational universals, and their values are often functionally motivated externally (Croft 2008: 230). However, typological features nevertheless offer a useful basis of comparison along predefined feature values, and are therefore most suitable for synchronic clustering/similarity measurements (Haspelmath 2010), such as the ones we are aiming for in our study.

11.6 Areal clusters

From the analysis of the 130 features from the 19 creole varieties in SplitsTree, we obtain the following graph. It is evident from Figure 11.1 that a signal of similarity can be established in these data. Languages that we know to be closely related from their external history, such as the Cape Verdean varieties or the Chabacano varieties, appear closer together in the graph. This means that languages that are closely related show similar grammatical and lexical patterns.

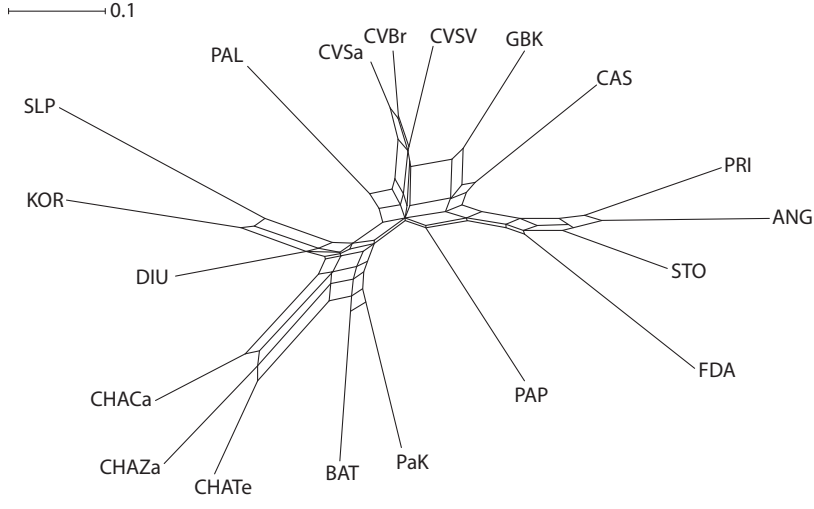


Figure 11.1 A split graph of a Neighbor-Net analysis of the 130 APiCS features in 19 Iberian creoles.

The general picture that can be observed from the graph is the division into Atlantic and Asian creoles. The split between these two main geographical areas is visible in the narrow area in the middle of the graph, with the Atlantic creoles on the right of the graph and the Asian creoles on the left. The main subclusters that emerge from the analysis are composed of the Gulf of Guinea creoles and Chabacano. Other, weaker clusters are formed by the Indo-Portuguese creoles, the Cape Verdean varieties, and the Malayo-Portuguese creoles. In addition, Guinea-Bissau Kriyol and Casamancese Creole appear between the Cape Verdean and the Gulf of Guinea clusters. However, there are also divergences from the pattern; Palenquero and Papiamentu show long independent lines on the Atlantic side, and thus do not form a Caribbean cluster.

The delta scores of the graph in Figure 11.1 are given in Table 11.6 for each of the 19 Iberian creoles. The average delta score is 0.3194. The lowest scores are among the languages that cluster together and are more separate from the webbing,

such as the Chabacano varieties and the Gulf of Guinea creoles. Languages that have a higher delta score are Palenquero, Papiamentu, and Guinea-Bissau Kriyol. In this particular case, it seems that these higher scores are connected to conflicting feature groups in the sample. This means that, for a part of the feature sample, these are similar to certain languages that form areal clusters, while for another part they pattern similarly to other languages in other areal groups. Therefore, the graph in Figure 11.1 should be interpreted in the first place as a map of typological types (Verkerk 2012: 73), which are based on areal features conditioned largely by substrate and adstrate influence.

Table 11.6 Delta scores for the Neighbor-Net analysis of 130 APiCS features in 19 Iberian creoles.

Language	Delta score
Zamboanga Chabacano	0.26162
Ternate Chabacano	0.26823
Principense	0.28064
Angolar	0.28412
Cape Verdean of Santiago	0.29853
Korlai	0.29901
Cavite Chabacano	0.29988
Santome	0.30374
Cape Verdean of Brava	0.31481
Sri Lanka Portuguese	0.32276
Fa d'Ambô	0.33137
Cape Verdean of São Vicente	0.33955
Diu	0.34191
Casamancese	0.34275
Papiá Kristang	0.34866
Batavia	0.35010
Papiamentu	0.35586
Guinea-Bissau Kriyol	0.35616
Palenquero	0.36984

Further testing was also done by selecting only one variety from each regional area and bi-clan (see Michaelis 2014), i.e. a combination of a lexifier clan and a substrate clan, but the results were very similar, although the number of languages was often very small. The Atlantic-Asian divide stays clear, as does the division into South Asian and Southeast Asian creoles, albeit to a somewhat lesser extent. Independent lines are long for all the varieties showing divergence on the level of individual languages. However, among the Atlantic creoles, regional clusters are not clear. The inclusion of the lexifiers, Portuguese and Spanish, does not affect

the clustering in Figure 11.1. The lexifiers are placed in the middle of the Atlantic-Asian divide, and their inclusion increases the webbing on the Atlantic side of the graph. This suggests that the lexifier, which in this case is Portuguese, shares mixed feature sets with the different Atlantic creoles, and these shared features do not follow areal clusters. As the results are not significantly different, we have decided not to reproduce the graphs here.

Subsets of features were also analysed. Phonological features from the APiCS database produced a mixed picture, due to the low number of features (13). Similarly, lexico-semantic information from the APiCS database did not produce any relevant clustering, as the number of the features (9) was too small. Networks with 108 structural/functional features from the APiCS database gave results that reproduce the groupings in Figure 11.1. These were chosen by excluding the lexical and phonological features. In a similar vein, a selection of 21 features from the nominal domain (features 2–7 and 9–10, focusing on word order, and features 22–34 and 37–40, focusing on nominal categories and nominal syntax) gives very similar groupings, with the exception of clustering Cape Verdean creoles on the same side with Asian creoles, as shown in Figure 11.2. This is probably due to a larger number of features that the Asian creoles share with the lexifier in this domain. When it comes to the divided clustering of Upper Guinea creoles, the results can correspondingly be interpreted as stronger adstrate influence on the continental Upper Guinea creoles Casamancese creole and Guinea-Bissau Kriyol.

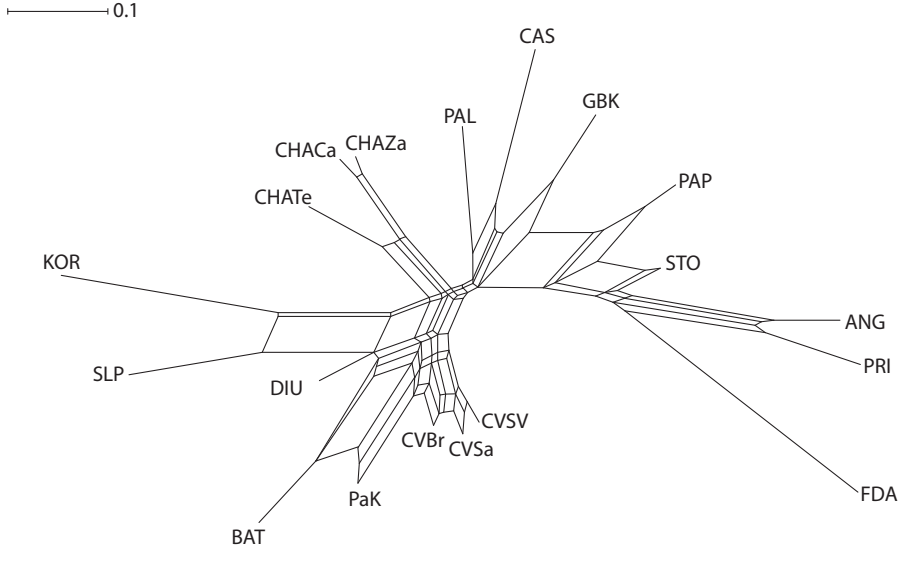


Figure 11.2 A splits graph of a Neighbor-Net analysis of 21 APiCS features of the nominal domain in 19 Iberian creoles.

11.7 Shared features

In addition to the Neighbor-Net analysis that reveals the areal clusters in our sample, we can conduct a more fine-grained analysis to reveal similarities and differences on the feature level in order to see which features are typical for each cluster. The features are mapped using parsimonious trees created in PAUP*, which are later analysed with Mesquite in order to detect cases of independent innovation or differences through the reconstruction of a given linguistic feature. Although we are not directly interested in the reconstruction of ancestral states, the tool is useful for detecting shared features in the established clusters.

The results of the analysis show shared features across different parts of grammar. We will go through them starting with features that are common to all the Iberian creoles in the sample, and thereafter we will present the features that are subject to some exceptions.

Table 11.7 presents the features that are shared by all the varieties in the sample. Iberian creoles present no dual pronouns, they have no syncretism in independent personal pronouns nor applicative constructions, and they mark instrumental expressions by adpositions.

Table 11.7 Shared features among 19 Iberian creoles.

No. in APiCS	Shared features	Feature category
14	No dual in independent personal pronouns	Nominal categories
16	No person syncretism in independent personal pronouns	
27	No antidual of paired body-part terms	
69	Instrumental expressions by adpositions	Argument marking
91	No applicative constructions	Clausal syntax

These features are also commonly shared among creoles in general. For example, applicative constructions are extremely rare among creoles and can only be found in the African contact varieties and the non-creole Michif (Haspelmath et al. 2013). Antidual of paired body-part terms is a feature found only among the French creoles of the Indian Ocean, Haitian Creole, Nengee, and Creolese, and not pervasively in any of them (Michaelis & the APiCS Consortium 2013a).

Two of the shared features pertain to independent personal pronouns. Dual personal pronouns are atypical of creoles in general, and this feature is observed only in Pichi in the Atlantic area, and in a number of creoles spoken in the Pacific and Australia (Haspelmath & the APiCS Consortium 2013a). Here, an observation should also be made about grammaticalization processes that are in progress, which the data included in typological databases do not depict. For example, in Ternate Chabacano, the expression of duality is possible by adding the numeral

dos 'two' to the pronoun, as in *mótru-dos* 'we.dual', although the construction is not fully grammaticalized. Another shared feature from the area of personal pronouns is the lack of person syncretism, reflecting yet another general tendency in creoles. This feature is typical only in the Indian Ocean French creoles in the second and third person. It is also observed in some other varieties, such as Haitian Creole and Nengee in the first and second person (Haspelmath & the APiCS Consortium 2013b).

In the case of the instrumental expressions, the picture is less clear for the Iberian varieties due to different optional strategies that are common among the Atlantic creoles. Several creoles in the Caribbean and in West Africa use serial verbs to express instrumental relations. Actually, both adpositional marking and serial verb strategies are possible in the Gulf of Guinea creoles, with the exception of Fa d'Ambô, and in some Upper Guinea creoles, such as Casamancese Creole and the Santiago and Brava varieties of Cape Verdean Creole (Maurer & the APiCS Consortium 2013).

An even larger number of features is shared if exceptions are permitted. Table 11.8 lists the features for which only one variety in our sample presents an exception. Differences are generally not representative of only one variety, although Sri Lanka Portuguese stands out in several cases.

Table 11.8 Features shared across 18 of the 19 Iberian creoles.

No. in APiCS	Shared features	Feature category
4	Order of adposition and noun phrase	Word order
6	Order of cardinal numeral and noun	
7	Order of relative clause and noun	
10	Position of indefinite article in the noun phrase	
15	No inclusive/exclusive distinction in independent personal pronouns	Nominal categories
36	No sortal numeral classifiers	
41	Comparative adjective marking	Nominal syntax
97	'Want' complement subjects left implicit	Complex sentences
99	Verb doubling not possible in temporal clauses	
110	Savvy word exists	Lexicon

Sri Lanka Portuguese is distinct in not allowing the same strategies as the other Iberian creoles for five features, for which we find parallels in the adstrate languages. Generally, Iberian creoles all have prepositions, and in all of them the relative clause can be placed after the noun. In Sri Lanka Portuguese, however, the order features are exceptional and do not follow the general tendencies. The adpositions are postpositions, and the relative clause precedes the noun (Smith

2013). Korlai Portuguese also uses postpositions to a considerable degree, but it allows prepositions as well, and has correlative clauses as the predominant relative clause type (Clements 2013). Also, contrary to the prevalent strategy among the Iberian creoles, Sri Lanka Portuguese does not mark the adjective in comparatives (Cardoso 2012: 110–111). This strategy is also possible in the Gulf of Guinea creoles, with the exception of Angolar, and in Ternate Chabacano. Among creoles in general, non-marking of the adjective in comparative constructions can be observed in African and Australian varieties (Michaelis & the APiCS Consortium 2013b). Another difference is that Sri Lanka Portuguese uses a desiderative particle in same-subject complement clauses of 'want', while most creoles use a verb in a construction where the complement subject is left implicit (Michaelis, Haspelmath & the APiCS Consortium 2013). Lastly, regarding numeral classifiers, which are not found in Iberian creoles, Sri Lanka Portuguese occasionally uses *pesaam* 'person' as a classifier for humans, reflecting a pattern found in the adstrate Tamil/Sinhala.

In addition, Zamboanga Chabacano differentiates between inclusive and exclusive personal pronouns, which is an adstrate feature. Principense presents postposed indefinite articles, which are pre-posed in other Iberian creole varieties. The Santiago variety of Cape Verdean Creole allows verb doubling constructions to be used to express temporal adverbial clauses. Lastly, Angolar does not have a word like *save* or *sabi* meaning 'know' (or something similar), deriving ultimately from forms of Portuguese or Spanish *saber*, which can be found in all other Iberian creoles.

The list of shared features that are found in the majority of the Iberian creoles is even longer, but variation comes increasingly into play when looking at other features. Table 11.9 presents the features that are shared by most of the varieties, but which nevertheless present two or three exceptions. Clear cases of exceptions are the following: Fa d'Ambô, Guinea-Bissau Kriyol, and Casamancese Creole have an indefinite article distinct from the numeral 'one' (feature 29); Principense and Santome verbal particles follow a Mood-Tense-Aspect order, but this feature does not apply for other Iberian creoles (feature 44); and Fa d'Ambô and Papiamentu have a neutral alignment in the case marking of personal pronouns (feature 59), whereas the rest of the Iberian creoles has an accusative alignment.

Many creoles permit several strategies for expressing a function, which makes the picture somewhat more complicated. For example, Fa d'Ambô is the only creole in the sample that only allows degree words following adjectives (feature 8), but this strategy is also found as a parallel to the shared pattern found in several other creoles in different geographical areas. These include the other Gulf of Guinea creoles, Casamancese Creole, Guinea-Bissau Kriyol, Cape Verdean Creole of Santiago, Korlai, Papiá Kristang, and Palenquero. On the other hand, Fa d'Ambô and Angolar do not mark the possessor noun phrase in an adnominal possessive construction (feature 38). Other varieties, such as Principense, Palenquero, and

Table 11.9 Features that are shared by most of the Iberian creoles.

No.	Shared features	Feature category	Exceptions
8	Order of degree word and adjective	Word order	Fa d'Ambô
12	Position of interrogative phrases in content questions: initial		Diu, Korlai, Sri Lanka
20	Pronoun conjunction: singular pronoun overtly conjoined with other conjunct	Nominal categories	Angolar, Chabacano
29	Indefinite articles identical to 'one'		Fa d'Ambô, Guinea-Bissau Kriyol, Casamancese Creole
38	Adpositional or case marking of possessor	Nominal syntax	Fa d'Ambô, Angolar
44	Order of tense, aspect, and mood markers does not apply	Verbal categories	Santome, Principense
59	Accusative alignment of case marking of personal pronouns	Argument marking	Fa d'Ambô, Papiamentu
67	Experiencer in subject position in constructions with 'like'		Cape Verdean Creole of São Vicente, Casamancese Creole, Korlai, Sri Lanka
81	Motion-to and motion-from differentiated	Clausal syntax	Cavite and Zamboanga Chabacano

Cape Verdean Creole of Santiago, have this strategy as an option as well, but the creoles that do not belong to the Gulf of Guinea cluster tend to use adpositions or case to mark the possessor. The latter strategy is typical for the rest of the Iberian creoles. Similarly, although interrogative phrases in content questions are generally initial among Iberian creoles (feature 12), the Indo-Portuguese varieties present a clear exception, and a number of other Portuguese-lexifier creoles also permit the placement elsewhere.

Generally, Iberian creoles overtly conjoin singular personal pronouns and full noun phrases, as in 'Mary and I', or 'you and your brother' (feature 20). Only Angolar does not permit this strategy, and variation is also found in Cavite and Ternate Chabacano. Differentiation between motion-to and motion-from (feature 81), on the other hand, presents a clear case of variation that is reflected in the Chabacano lects described in APiCS, where the values of identity (with *na*), overlap (*na*, *para na*), and differentiation (*na*, *di*) are assigned to different varieties.

An overview of the shared features and their comparison with areal clusters and lexifier features reveals that most of the shared varieties are also shared with the Portuguese and Spanish lexifiers. A clearly distinct feature that differentiates the creole varieties from the lexifiers, but presents a lot of variation among the creoles,

is the position of TMA marking (feature 43). It is a well-known fact that creoles are set apart from their Spanish and Portuguese lexifiers due to the differences in the verbal system. However, this feature is also a nearly universal creole trait.

11.8 Areal differences

The features that define the areal clusters in Figure 11.1 can be identified in Mesquite. The same areal clusters can also be identified in the PAUP analysis that creates the parsimonious tree in Figure 11.3. The lexifiers stay apart from the creoles, and the main geographical division between Atlantic and Asian is clear.

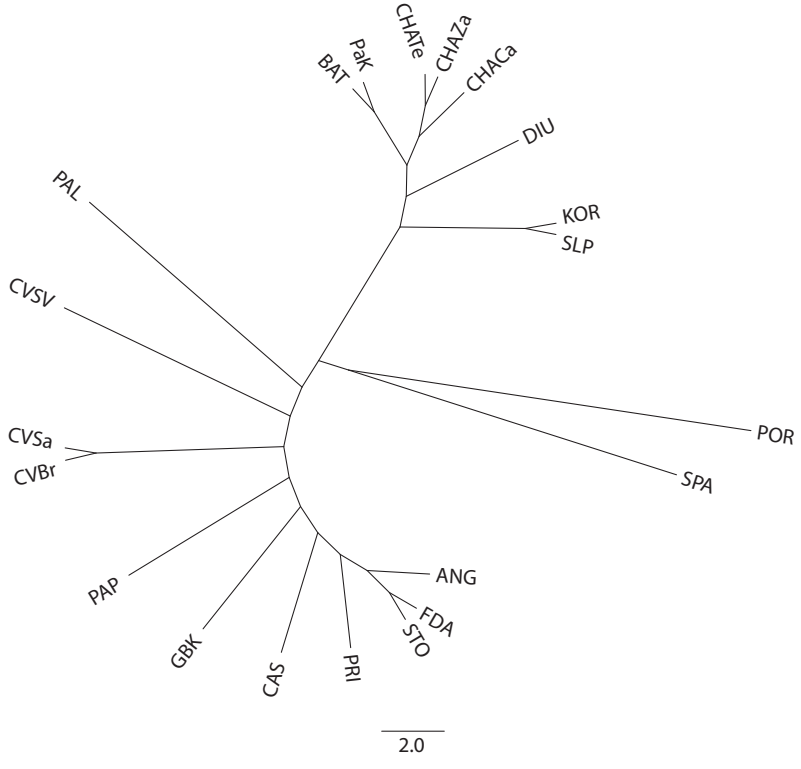


Figure 11.3 Unrooted Maximum Parsimony tree of the Iberian creoles.

However, only two clusters get sufficient statistical support: the Gulf of Guinea creoles and the Chabacano varieties of the Philippines. Tables 11.10 and 11.11 list the features (with their corresponding APiCS feature number and the assigned values for these varieties) that appear only in these clusters and not in any other Iberian creoles in our sample.

Table 11.10 Gulf of Guinea features.

Gulf of Guinea creoles	
38	No marking of possessor noun phrases
80	Serial verb constructions to express 'motion-from'
82	Serial verb construction (without a preposition) to express the transitive motion construction with 'push'
83	Serial verb construction (without a preposition) to express the transitive motion construction with 'pull'
100	Bipartite negative marker
101	Bipartite negative marker, before verb and after object
107	Optional vocative marker following noun

The use of serial verb constructions to express motion is characteristic of the Gulf of Guinea cluster, and it appears only very occasionally in other contact languages of the area that have a non-Iberian lexifier, e.g. Ghanaian Pidgin English, Nigerian Pidgin, and Sango. No marking of possessor noun phrases is more common in the Atlantic creoles, whereas optional vocative markers following the noun seems to be an areal feature. Bipartite negative marking is also possible in Palenquero, although it is not the only strategy available. Hagemeijer (2011: 140) connects these features and some others to the Western Bantu (Kikongo) and Niger Delta adstrates.

Table 11.11 Chabacano features.

Chabacano/Philippines	
1	VSO
61	Verb-initial theme – recipient
102	Negation and indefinite pronouns

The features that are characteristic of the Chabacano creoles are related to verb-initial word order, which is a Philippine feature. Similarly, indefinite pronouns are formed by an existential construction, which is the most common strategy in the negative paradigm. This strategy also reflects adstrate structures (see also Sippola 2012 for a comparison of the variation in the Ibero-Asian creoles).

A look at the other features also shows some that are shared within areal groups. The Southeast Asian creoles, Chabacano, Papiá Kristang, and Batavia creole, share purely aspectual TMA systems (feature 49) and give-serials, where 'give' appears in first position expressing the recipient only (feature 86). The Indo-Portuguese varieties share postposed adpositions in NPs (feature 4), non-initial interrogative phrases (feature 12), and the possibility of using locational constructions to express predicative possession (feature 77). In continental Upper Guinea, the same marker is shared for the progressive and the future (feature 47).

11.9 Discussion

The results of our analysis confirm several areal clusters that were identified in previous research, such as the Cape Verdean varieties, the Gulf of Guinea creoles, Indo-Portuguese, Malayo-Portuguese, and the Chabacano varieties. Other Upper Guinea creoles appear more independently. That is also the case for the Caribbean creoles Papiamentu and Palenquero. This can be expected, as they are the only ones in South America and the Caribbean. Regarding the method, it is interesting to see that both the Neighbor-Net and the maximum parsimony analysis produced very similar results. When comparing these results with earlier classifications, we notice that the same lines of groupings have been established with the help of individual features or sub-selections of them in previous research. When comparing, for example, Ferraz's (1987) features in Table 11.1 to the APiCS data, it is clear that the features partially match. Some features from Ferraz's analysis, such as the diminutives, the particle of obligation, the unmarked personal pronoun for the Atlantic creoles, the – *mente* adverbials, and the forms of the TMA and object markers for Asian creoles are missing from the APiCS data, in addition to the lexical information. However, the features in APiCS include a much wider range of structural and functional information. Clements' (2009) affixal TMA marker and grammatical features are also present in the APiCS data, but they did not appear as decisive for individual groupings.

The shared core of the Iberian creoles is a combination of typical creole traits and lexifier influence, as the naming of this group suggests. The features are to a great extent shared with the lexifiers. However, previous features, mostly based on the form of lexical and functional items, were not included in the APiCS sample (for example, the features in Holm 2009). Despite this, the number of shared features is considerable at 26 out of 130, when two to three exceptions are permitted. It will be interesting to see how different lexifier groups behave in this respect. When the shared features are divided according to their categories, word order and nominal categories are found to be salient, as can be seen in Table 11.11. As the original division into different categories of grammar was not even, we did further testing with regression analysis to see if the overall number of features in each category predicts the number of shared features in a meaningful manner. This was not the case, as a 1% change in the original number of features in a category leads to a 1.1% increase in the shared features in the same category. However, this might be due to the fact that we have a small number of data points.

In a recent phylogenetic study on the typological status of creole languages, Bakker et al. (2011) used 97 features selected in *Comparative Creole Syntax* (Holm & Patrick 2007) to compare creoles from different lexifier and areal groups, including seven Iberian creoles. Their results (Bakker et al. 2011: 29) did not provide

Table 11.12 Shared features according to the feature categories.

Category	Shared features (max. 3 dev.) / all	% of shared features	Shared in the feature category
Nominal categories	7 / 24	26.9%	29.2%
Clausal syntax	2 / 19	7.7%	10.6%
Argument marking	3 / 16	11.5%	18.8%
Verbal categories	2 / 14	7.7%	14.3%
Phonology	0 / 13	0.0%	0%
Word order	6 / 12	23.1%	50%
Lexicon	1 / 11	3.8%	9.1%
Complex sentences	2 / 8	7.7%	25.0%
Negation, questions, focusing	1 / 7	3.8%	14.3%
Nominal syntax	2 / 6	7.7%	33.3%
Total	26 / 130	100%	

evidence for claims about the decisiveness of lexifier or substrate influence for the groupings of the creoles. Although there was some partial clustering according to the lexifier among the seven Iberian creoles compared, when contrasted with creoles with a different lexifier, the clustering was not clear. The only Iberian varieties that were clearly connected in their network analysis were Cape Verdean and Guinea-Bissau Kriyol. In our data, the clustering of Cape Verdean and Guinea-Bissau Kriyol is not as evident, as seen in Figure 11.1. While their study was the first one to apply phylogenetic methods to a creole sample including Iberian varieties, it did not identify the specific features that characterize different clusters. In the section focusing on the differences, we have identified areal features which are also found in the substrate and adstrate languages of these creoles, showing that these do play a role.

Another point to keep in mind is that the sample as a whole has not gone through an assessment of areality in the typological sense (Bisang 2010: 419). This means that the results shed light on the typical features of this subgroup, which might or might not correspond to those of the creoles (see Bakker & Daval-Markussen this volume, chapter 6). Geography or areality is here understood as "individual sets of linguistic features which are pervasively observable in a particular world region or a part thereof against all others" (Kortmann 2013: 166). It is clear, however, that several features in the proposed areal clusters are not independent of one another, as shown for example by the fact that several features were about the use of serial verbs in motion constructions in the Gulf of Guinea cluster, and about the verb-initial order in the Philippines.

11.10 Conclusions

Computational analysis of typological data supports previous classifications of Iberian creoles that were established in early creole studies and refined in subsequent comparative work. The results show a clear division into Atlantic and Asian creoles, and the main subgroups were divided geographically. However, connections between the subgroups are less clear due to the degree of detail of the typological feature values, as clusters reflect mainly areal traits in these languages. The synchronic data used in this study thus provides no support for theories about the connections between Papiamentu and the Upper Guinea creoles, or the interrelations between the Chabacano varieties and the Portuguese creoles of Asia.

A mapping of the features responsible for this classification confirms that many similarities among the Iberian creoles are shared with the lexifiers. They also present universal tendencies that are further conditioned by the individual contact situations. Differences and characteristic traits of the Gulf of Guinea and Chabacano creoles, as well as Indo-Portuguese and the continental Upper Guinea creoles, are clearly defined by the adstrate and substrate languages, and thus lead to the areal patterning.

Notes

The datasets for this chapter can be found here: https://phylogenetic-creole-studies.blogspot.com

References

Bakker, P., Daval-Markussen, A., Parkvall, M. & Plag, I. 2011. Creoles are typologically distinct from non-creoles. *Journal of Pidgin and Creole Languages* 26(1): 5–42. doi: 10.1075/jpcl.26.1.02bak

Baptista, M., Mello, H. & Suzuki, M. 2007. The syntax of Cape Verdean Creole and Guinea-Bissau Creole. In *Comparative Creole Syntax. Parallel Outlines of 18 Creole Grammars* [Westminster Creolistics 7], J. Holm & P. Patrick (eds), 53–82. London: Battlebridge.

Batalha, G. N. 1988. *Suplemento ao glossário do dialecto macaense: novas notas linguísticas, etnográficas e folclóricas*. Macau: Instituto Cultural de Macau.

Baxter, Alan. 1996. Portuguese and Creole Portuguese in the Pacific and Western Pacific Rim. In *Atlas of Languages of Intercultural Communication in the Pacific, Asia and the Americas*, S. A. Wurm, P. Mühlhäusler & D. T. Tryon (eds), 299–338. Berlin: Mouton de Gruyter. doi: 10.1515/9783110819724.2.299

Bryant D. & Moulton V. 2004. Neighbor-net: An agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution* 21(2): 255–265. doi: 10.1093/molbev/msh018

Bisang, W. 2010. Areal language typology. In *Language and Space. An International Handbook of Linguistic Variation*, Vol. 1: *Theories and Methods* [Handbücher zur Sprach- und Kommunikationswissenschaft], P. Auer & J. E. Schmidt (eds), 419–440. Berlin: Mouton de Gruyter.

Cardoso, H. C. 2010. The African slave population of Portuguese India: Demographics and impact on Indo-Portuguese. *Journal of Pidgin and Creole Languages* 25(1): 95–119. doi: 10.1075/jpcl.25.1.04car

Cardoso, H., Baxter, A. N. & Pinharanda Nunes. M. (eds). 2012. *Ibero-Asian Creoles. Comparative Perspectives* [Creole Language Library 46]. Amsterdam: John Benjamins. doi: 10.1075/cll.46

Cardoso, H. 2012. Luso-Asian comparatives in comparison. In *Ibero-Asian Creoles: Comparative Perspectives* [Creole Language Library 46], H. C. Cardoso, A. N. Baxter & M. Pinharanda Nunes (eds), 81–124. Amsterdam: John Benjamins. doi: 10.1075/cll.46.04car

Clements, C. 2009. *The Linguistic Legacy of Spanish and Portuguese. Colonial Expansion and Language Change*. Cambridge: CUP. doi: 10.1017/CBO9780511576171

Clements, J. C. 2013. Korlai structure dataset. In *Atlas of Pidgin and Creole Language Structures Online*, S. M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/contributions/40> (25 September 2014).

Croft, W. 1991. *Typology and Universals*. Cambridge: CUP.

Croft, W. 2008. Evolutionary linguistics. *Annual Review of Anthropology* 37: 219–34. doi: 10.1146/annurev.anthro.37.081407.085156

Coelho, F. A. 1881. *Os dialectos romanicos ou neo-latinos na África, Ásia e América*. Lisboa: Casa da Soc. de Geografia.

Dalgado, S. R. 1921. *Glossário luso-asiático*. Coimbra: Imprensa da Universidade.

Daval-Markussen, A. & Bakker, P. 2011. A phylogenetic networks approach to the classification of English-based Atlantic Creoles. *English World-Wide* 32(2): 115–146. doi: 10.1075/eww.32.2.01dav

Daval-Markussen, A. & Bakker, P. 2012. Explorations in creole research with phylogenetic tools. In *Proceedings of the EACL 2012 Joint Workshop of LINGVIS & UNCLH*, 89–97. Avignon, France, April 2012. Association for Computational Linguistics.

Dunn, M. 2009. Contact and phylogeny in Island Melanesia. *Lingua* 119: 1664–1678. doi: 10.1016/j.lingua.2007.10.026

Fernández, M. A. 2004. Plurifuncionalidad de la partícula na en el chabacano de Zamboanga. In *Los criollos de base ibérica. ACBLPE 2003*, M. Fernández, M. Fernández-Ferreiral & N. Vázquez Veiga (eds), 41–59. Frankfurt & Madrid: Vervuert & Iberoamericana.

Fernández, M. A. 2007. Sobre el origen de *con* en chabacano. In *La Romania en interacción: Entre historia, contacto y política. Ensayos en homenaje a Klaus Zimmermann*, M. Schrader-Kniffki & L. Morgenthaer García (eds), 457–478. Frankfurt & Madrid: Vervuert & Iberoamericana.

Fernández, M. A. 2012. Leyenda e historia del chabacano de Ermita (Manila). *UniverSOS* 9: 9–46.

Ferraz, L. I. 1979. *The Creole of São Tomé*. Johannesburg: Witwatersrand University Press.

Ferraz, L. I. 1987. Portuguese creoles of West Africa and Asia. In *Pidgin and Creole Languages: Essays in Memory of John E. Reinecke*, G. G. Gilbert (ed.), 337–360. Honolulu HI: University of Hawai'i Press.

Gray, R. D., Bryant, D. & Greenhill, S. J. 2010. On the shape and fabric of human history. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365(1559): 3923–3933. doi: 10.1098/rstb.2010.0162

Hancock, I. F. 1975. Malacca Creole Portuguese: Asian, African or European? *Anthropological Linguistics* 17: 211–236.

Hagemeijer, T. 2011. The Gulf of Guinea creoles: Genetic and typological relations. *Journal of Pidgin and Creole Languages* 26(1): 111–154. doi: 10.1075/jpcl.26.1.05hag

Haspelmath, M. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3): 663–687. doi: 10.1353/lan.2010.0021

Haspelmath, M. & the APiCS Consortium. 2013a. Dual in independent personal pronouns. In *Atlas of Pidgin and Creole Language Structures Online*, S. M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/parameters/14> (25 September 2014).

Haspelmath, M. & the APiCS Consortium. 2013b. Person syncretism in independent personal pronouns. In *Atlas of Pidgin and Creole Language Structures Online*, S. M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/parameters/16> (24 September 2014).

Haspelmath, M., Michaelis, S. M. & the APiCS Consortium. 2013. Applicative constructions. In *Atlas of Pidgin and Creole Language Structures Online*, S. M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/parameters/91> (25 September 2014).

Holm, J. 1988–1989. *Pidgins and Creoles*. Cambridge: CUP.

Holm, J. 2009. Atlantic features in Asian varieties of Creole Portuguese. *Journal of Portuguese Linguistics* 8(2): 11–22.

Holm, J. & Patrick, P. 2007. *Comparative Creole Syntax, Parallel outlines of 18 creole grammars* [Westminster Creolistics 7]. London: Battlebridge.

Huson, D. H. & Bryant, D. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23(2): 254–267. doi: 10.1093/molbev/msj030

Jacobs, B. 2009. The Upper Guinea origins of Papiamentu. Linguistic and historical evidence. *Diachronica* 26(3): 319–379. doi: 10.1075/dia.26.3.02jac

Jacobs, B. 2012. *Origins of a Creole. The History of Papiamentu and its African Ties*. Berlin: Mouton. doi: 10.1515/9781614511076

Kihm, A. & Rouge, J.-L. 2013. Língua de Preto, the basic variety at the root of West African Portuguese Creoles. A contribution to the theory of pidgin/creole formation as second language acquisition. *Journal of Pidgin and Creole Languages* 28(2): 203–298. doi: 10.1075/jpcl.28.2.01kih

Kortmann, B. 2013. How powerful is geography as an explanatory factor of variation? Areal features in the Anglophone world. In *Space in Language and Linguistics: Geographical, Interactional, and Cognitive Perspectives*, P. Auer, M. Hilpert, A. Stukenbrock & B. Szmrecsanyi (eds), 165–194. Berlin: Walter de Gruyter. doi: 10.1515/9783110312027.165

Lipski, J. 1988. Philippine Creole Spanish: Assessing the Portuguese element. *Zeitschrift für Romanische Philologie* 104: 25–45. doi: 10.1515/zrph.1988.104.1-2.25

Lipski, J. 1992. New thoughts on the origins of Zamboangueño (Philippine Creole Spanish). *Language Sciences* 14(3): 197–231. doi: 10.1016/0388-0001(92)90005-Y

Lipski, J. M. 2008. Spanish-based creoles in the Caribbean. In *The Handbook of Pidgin and Creole Studies*, S. Kouwenberg & J. V. Singler (eds), 543–564. Oxford: Wiley-Blackwell.

List, J-M., Nelson-Sathi, S., Geisler, H. & Martin, W. 2013. Networks of lexical borrowing and lateral gene transfer in language and genome evolution. *Bioessays* 36: 141–150. doi: 10.1002/bies.201300096

Maddison, W. P. & Maddison, D. R. 2014. Mesquite: A modular system for evolutionary analysis. Version 3.0 <http://mesquiteproject.org> (13 May 2015).

Maurer, P. & the APiCS Consortium. 2013. Instrumental expressions. In *Atlas of Pidgin and Creole Language Structures Online*, S. M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/parameters/69> (26 September 2014).

Michaelis, S. M. 2014. Sampling in contact linguistics: What is a typical creole feature? Paper presented at *Time and Space in Linguistics: Interdisciplinary Computational Approaches & Cross-Creole Comparisons*, Aarhus, 15–16 January.

Michaelis, S. M. & the APiCS Consortium. 2013a. Antidual of paired body-part terms. In *Atlas of Pidgin and Creole Language Structures Online*, S. M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/parameters/27> (25 September 2014).

Michaelis, S. M. & the APiCS Consortium. 2013b. Comparative adjective marking. In *Atlas of Pidgin and Creole Language Structures Online*, S. M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/parameters/41> (27 September 2014).

Michaelis, S. M., Haspelmath, M. & the APiCS Consortium. 2013. 'Want' complement subjects. In *Atlas of Pidgin and Creole Language Structures Online*, S. M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/parameters/97> (25 September 2014).

Michaelis, S. M., Maurer, P., Haspelmath, M. & Huber, M. (eds). 2013. *Atlas of Pidgin and Creole Language Structures Online*, Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info> (25 September 2014).

Naro, A. J. 1978. A study on the origins of pidginization. *Language* 54(2): 314–347. doi: 10.2307/412950

Parkvall, M. 2000. *Out of Africa. African Influences in Atlantic Creole Languages*. London: Battlebridge.

Quint, N. 2000. *Le Cap-Verdien. Origines et devenir d'une langue métisse (Etude des relations de la langue cap-verdienne avec les langues africaines, créoles et portugaise)*. Paris: L'Harmattan.

Schuchardt, H. 1882a. Kreolische Studien I. Über das Negerportugiesische von S. Thomé (Westafrika). *Sitzungsberichte der Philosophisch-Historischen Classe der kaiserlichen Akademie der Wissenschaften* 101(2): 889–917.

Schuchardt, H. 1882b. Kreolische Studien II. Über das Indoportugiesische von Cochim. *Sitzungsberichte der Philosophisch-Historischen Classe der kaiserlichen Akademie der Wissenschaften* 102(2): 799–816.

Schuchardt, H. 1883a. Kreolische Studien III. Über das Indoportugiesische von Diu. *Sitzungsberichte der Philosophisch-Historischen Classe der kaiserlichen Akademie der Wissenschaften* 103(1): 799–816.

Schuchardt, H. 1883b. Kreolische Studien IV. Über das Malaiospanische der Philippinen. *Sitzungsberichte der Philosophisch-Historischen Classe der kaiserlichen Akademie der Wissenschaften* 105(1): 111–150.

Schuchardt, H. 1884. Kreolische Studien VI. Über das Indoportugiesische von Mangalore. *Sitzungsberichte der Philosophisch-Historischen Classe der kaiserlichen Akademie der Wissenschaften* 105(3): 881–904.

Schuchardt, H. 1888a. Beiträge zur Kenntnis des kreolischen Romanisch: I. Allgemeineres über das Negerportugiesische. *Zeitschrift für romanische Philologie* 12: 242–254. doi: 10.1515/zrph.1888.12.1-4.242

Schuchardt, H. 1888b. Beiträge zur Kenntnis des kreolischen Romanisch: II. Zum Negerportugiesischen Senegambiens. *Zeitschrift für romanische Philologie* 12: 301–312. doi: 10.1515/zrph.1888.12.1-4.301

Schuchardt, H. 1888c. Beiträge zur Kenntnis des kreolischen Romanisch: III. Zum Negerportugiesischen der Kapverden. *Zeitschrift für romanische Philologie* 12: 312–322. doi: 10.1515/zrph.1888.12.1-4.242

Schuchardt, H. 1888d. Kreolische Studien VII. Über das Negerportugiesische von Annobom. *Sitzungsberichte der Philosophisch-Historischen Classe der kaiserlichen Akademie der Wissenschaften* 116(1): 193–226.

Schuchardt, H. 1889a. Beiträge zur Kenntnis des kreolischen Romanisch: IV. Zum Negerportugiesischen der Ilha do Principe. *Zeitschrift für Romanische Philologie* 13: 463–475. doi: 10.1515/zrph.1889.13.1-4.463

Schuchardt, H. 1889b. Beiträge zur Kenntnis des kreolischen Romanisch: V. Allgemeineres über das Indoportugiesische (Asioportugiesische). *Zeitschrift für romanische Philologie* 13: 476–516. doi: 10.1515/zrph.1889.13.1-4.463

Schuchardt, H. 1889c. Beiträge zur Kenntnis des kreolischen Romanisch: VI. Zum Indoportugiesischen von Mahé und Cannanore. *Zeitschrift für romanische Philologie* 13: 516–524. doi: 10.1515/zrph.1889.13.1-4.463

Schuchardt, H. 1890. Kreolische Studien VII. Über das Malaioportugiesische von Batavia und Tugu. *Sitzungsberichte der Philosophisch-Historischen Classe der kaiserlichen Akademie der Wissenschaften* 122(9): 1–256.

Schwegler, A. & Green, K. 2007. Palenquero (Creole Spanish). In *Comparative Creole Syntax. Parallel Outlines of 18 Creole Grammars* [Westminster Creolistics 7], J. Holm & P. Patrick (eds), 273–306. London: Battlebridge.

Smith, I. R. 2013. Sri Lanka Portuguese structure dataset. In *Atlas of Pidgin and Creole Language Structures Online*, S. M. Michaelis, P. Maurer, M. Haspelmath & M. Huber (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/contributions/41> (25 September 2014).

Sippola, E. 2012. Indefinite terms in Ibero-Asian Creoles. In *Ibero-Asian Creoles: Comparative Perspectives* [Creole Language Library 46], H. C. Cardoso, A. N. Baxter & M. Pinharanda Nunes (eds), 149–180. Amsterdam: John Benjamins. doi: 10.1075/cll.46.06sip

Steinkrüger, P. O. 2006. The puzzling case of Chabacano: Creolization, substrate, mixing and secondary contact. Paper presented at Tenth International Conference on Austronesian Linguistics, Puerto Princesa City, Palawan, Philippines, 17–20 January. <http: //www.sil.org/asia/philippines/ical/papers.html> (5 September 2014).

Swofford, DL. 2002. PAUP*: Phylogenetic analysis using parsimony (* and other methods). Version 4.0b10. Sunderland MA: Sinauer Associates.

Thompson, R. W. 1961. A note on some possible affinities between the creole dialects of the Old World and those of the New. *Creole Language Studies* 2: 107–113.

Taylor, D. 1971. Grammatical and lexical affinities of creoles. In Pidginization and Creolization of Languages. Proceedings of a conference held at the University of the West Indies, Mona Jamaica, April 1968, D. Hymes (ed.), 293–296. Cambridge: CUP.

Vázquez Veiga, N. & Fernández, M. A. 2012. Maskin, maski, masque… in the Spanish and Portuguese creoles of Asia. In *Ibero-Asian Creoles: Comparative Perspectives* [Creole Language Library 46], H. C. Cardoso, A. N. Baxter & M. Pinharanda Nunes (eds), 181–204. Amsterdam: John Benjamins. doi: 10.1075/cll.46.07vaz

Verkerk, A. 2014. The Evolutionary Dynamics of Motion Event Encoding. PhD dissertation, Radboud University Nijmegen.

Wichmann S., Holman, E. W., Rama T. & Walker, R. S. 2011. Correlates of reticulation in linguistic phylogenies. *Language Dynamics and Change* 1: 205–240. doi: 10.1163/221058212X648072

Whinnom, K. 1956. *Spanish Contact Vernaculars in the Philippine Islands*. Hong Kong: Hong Kong University Press.

Whinnom, K. 1965. The origin of the European-based creoles and pidgins. *Orbis* 14: 509–527.