

**Bare Nouns among and beyond Creoles.**  
**A syntactic-semantic study of Kriyol Bare Noun Phrases**  
**based on a crosslinguistic comparison**  
**and the theoretical implications.**

**Dissertation**

zur Erlangung des akademischen Grades

**Dr. phil.**

**im Fach Allgemeine Sprachwissenschaft**

**von**

**Chiara Truppi**

Präsident der Humboldt Universität zu  
Berlin Prof. Dr. Jan-Hendrik Olbertz

Dekanin der Philosophischen Fakultät II  
Prof. Dr. Helga Schwalm

Gutachter:

1. Prof. Dr. Manfred Krifka - HU Berlin
2. Prof. Dr. Alain Kihm - CNRS Paris

eingereicht am 09.05.2014

verteidigt am 26.09.2014



# Acknowledgement

First of all, I would like to thank my advisor, Manfred Krifka, who has provided guidance and expertise; moreover, without his trust in my research project, I would not have had the possibility of a doctoral experience at the HU. Big thanks to Tonjes Veenstra for his support and his helpful suggestions during my years as a PhD student. I also want to thank Alain Kihm, Tjerk Hagemeijer, Elisabeth Verhoeven and Elena Gorishneva for having agreed to be part of the committee.

The Humboldt-Universität zu Berlin is a stimulating place to conduct research in semantics and syntax. I am grateful to the Studienstiftung des deutschen Volkes for the ideological and financial support given to me from August 2011 to July 2014.

Thanks to Ines Duarte, Tjerk Hagemeijer, Nélia Alexandre and Fernanda Pratas for offering me the possibility of being a guest PhD student at the CLUL of the Universidade de Lisboa. I also want to thank my friends and colleagues from the CLUL for their exchange and support: Eunice Fernandes, Paula Luegi, Patricia Nunes Gonçalves, João Silva, Francisco Costa, João Rodrigues, Celeste Rodrigues, Cacao (Maria do Carmo), Haldane Amaro and Abigail Cosme. Big thanks to John Holm for his support and interest in my work.

A special thanks to my colleagues from the GDRI – Seepicla research group on pidgin and creole languages for their precious feedback in our meetings in Berlin (June 2013) and in London (June 2014) and, in particular, to Viviane Depréz for her feedback and interest in my research. Thanks also to the audience of ACBLPE and SPCL Joint Meeting (June 2013, Lisbon) and FACS-4 (November 2014, Paris) for the precious feedback.

I also want to thank Susann Fischer and the scientific staff of the Department for Romance Linguistics of Hamburg University for having given me the possibility of presenting my research project at their doctoral colloquium, and for their feedback and interest in my work.

A special thanks to the research fellows of the ZAS Berlin for allowing me to present my research project in their DocSkills doctoral colloquium and for their precious feedback.

Among the colleagues that I want to thank for support there are Nuno Soares for his help and feedback on bare nouns in European Portuguese, and Michael Dürr, Katharina Böhnert, Mia Batinic, Antonio

Civardi, Marta Ghilardi, Shrita Hassamal, Maria Mazzoli, Feresteh Modarresi, Daniela Moncalvo, Kilu von Prinze, Katja Reetz, Cristina Villari, and Slawomir Wacewicz.

A very big thanks to my informants and friends in Lisbon, Berlin and Rome for their crucial help. Many of them allowed me to interview them, others were always ready to help me with any question relative to Kriyol and Guinea-Bissau: Mamadú Saibana Baldé, Michel Té, Intunda Na Montche, Armenio Mendes, Karina Silva Gomez, Jorge Sanca and the family Djalo, Candinha Pinto Maria, Elson Cabral, Edson Incopté, Francisco Mango, Junilto Alvarenga Ntchemó, Gerónimo Costa, Candinha Maria Pinto, Duka, Zuzu, Cambraima Alanso Cassama, Florinda Justino Silva, Ramatulai Djaló and Rashid Djaló. I apologize if anyone is missing.

A special thanks to my friends and colleagues of *Nô lanta djunto e.V.*, in particular to Pondingo Saliha von Medem and Leoni Sanca for their support.

Finally, my heartfelt thanks to Claudio and Carmen, Paolo and Roberta, and the rest of my family, and to my great friends Chiara Giuliante, Cristina Pareschi, Andrea Locatelli, Andrea Piccoli, Renzo Palazzetti, Antje Glaner, Liila Taruffi, Sven Stocks, Roberta Pozzone, Silvia Balassone, Sonia Bergamo, Ana Senčić, Charlene Neubert, Esther Doumbouya, and all the friends who have supported my interest and passion in linguistics. Thanks to Ezra Plessing for his support during the last steps of this experience. Also thanks to Adam Jaffee for reviewing my dissertation.

I apologize for anyone whom I did not mention here.

## **Abstract**

The nature of the present dissertation is threefold: i) descriptive, ii) comparative, and iii) theoretical. After a brief general discussion on the history and grammar of Guinea-Bissau Creole, and after an extensive review of various approaches on BNPs, both from the semantic and syntactic perspective, the present work will offer an exhaustive description of the distribution and interpretation of Bare Noun Phrases in GBC. They may be found in both argument and nonargument positions. The general tendency for BNPs in GBC is to yield a definite reading (subjects, recipient objects, in topicalization, dislocation and clefting). One difference is that bare patient objects may yield any possible interpretation, except from the specific plural.

BNPs interpretation is driven by contextual factors as well as by aspect and predicate type. Perfective and continuous imperfective contexts trigger definite specific readings for bare objects. One difference is that bare objects in habitual imperfective contexts yield indefinite nonspecific interpretations. As for predicate types, bare subjects of stage-level predicates yield existential readings, whereas bare subjects of individual-level predicates derive definite generic readings.

The present work also undertakes a crosslinguistic comparison between creoles and noncreoles: i) Cape Verdean Creole, Santome, Papiamentu and Brazilian Portuguese; and ii) Mandarin Chinese, Vietnamese and Gbe languages. It turns out that BNPs distribution and interpretation are quite homogeneous. Importantly, BNPs in any of these languages may yield both singular and plural readings: BNPs are thus unspecified as for Number. This leads us to our theoretical discussion on Number: starting from Depréz's (2007) Plural Parameter and its basic assumptions (e.g. BNPs are unspecified as for Number, and the basic denotation of nouns is kind of type e), a new model, and the consequent linguistic typology, is developed.

# Contents

List of abbreviations .....	x
Chapter 1.....	1
Introduction .....	1
1.1 Goals of the dissertation.....	1
1.2 Methodology.....	3
1.3 Outline of the dissertation.....	3
Chapter 2.....	5
A Sketch of Kriyol: Historical, social and linguistic aspects .....	5
2.1 Lusofonia .....	5
2.2 The rising of a new typology of languages: The creoles .....	6
2.2.1 Theories of creolization .....	9
2.3 The emergence of the Creole language of Guinea-Bissau.....	13
2.3.1 Kriyol in more recent times.....	15
2.3.2 Kriyol's <i>stratum</i> .....	18
2.4 Kriyol grammar.....	23
2.4.1 The morphosyntax of Kriyol.....	24
2.4.2 The nominal domain.....	28
2.4.3 The verbal domain .....	29
Chapter 3.....	36
Theoretical background .....	36

<b>3.1</b>	<b>Definiteness .....</b>	<b>36</b>
3.1.1	Approaches to definiteness.....	38
<b>3.2</b>	<b>Genericity.....</b>	<b>41</b>
3.2.1	Generic sentences.....	43
3.2.2	What makes a sentence (non)generic.....	43
3.2.3	Generic nouns .....	45
3.2.4	Taxonomic kind-denoting noun phrases .....	47
3.2.5	Nontaxonomic kind-referring noun phrases.....	49
<b>3.3</b>	<b>Specificity.....</b>	<b>50</b>
<b>3.4</b>	<b>Classification of sentence types .....</b>	<b>54</b>
<b>3.5</b>	<b>Carlson's (1977) theory .....</b>	<b>56</b>
3.5.1	Some problems faced by the Carlsonian approach .....	61
<b>3.6</b>	<b>Categories and types: On the interpretation of noun phrases.....</b>	<b>63</b>
3.6.1	Type-shifting operations: The inventory .....	64
3.6.2	Realization of type-shifting operations: Some examples.....	67
<b>3.7</b>	<b>Chierchia's Nominal Mapping Parameter .....</b>	<b>69</b>
<b>3.8</b>	<b>Longobardi's (1994) proposal.....</b>	<b>72</b>
3.8.1	A sketch of the Italian nominal system .....	72
3.8.2	The English case .....	74
3.8.3	Longobardi's parametric proposal .....	78
3.8.4	Expletive articles.....	80
3.8.5	Typological evidence .....	82
<b>3.9</b>	<b>Depréz's Plural Parameter .....</b>	<b>84</b>
<b>3.10</b>	<b>Bare nouns in European Portuguese: A general sketch .....</b>	<b>86</b>

<b>Chapter 4.....</b>	<b>92</b>
<b>The nominal system of Kriyol and its bare nouns.....</b>	<b>92</b>
<b>4.1 Bare noun phrases and the nominal system of Kriyol: An overview .....</b>	<b>92</b>
4.1.1 Definition of Bare Noun Phrases.....	93
<b>4.2 Kriyol nominal system .....</b>	<b>95</b>
4.2.1 Possible NPs in Kriyol .....	97
4.2.2 The Noun Phrase .....	99
4.2.3 The loss of ‘gender’ in the process of creolization.....	101
4.2.4 Number requirements .....	102
4.2.5 Mass and count Ns .....	104
4.2.6 Plurality in Kriyol: A tendency.....	105
4.2.7 More on animacy.....	110
4.2.8 The absence of the definite determiner .....	112
4.2.9 The indefinite determiner <i>un</i> .....	114
4.2.10 Demonstratives.....	118
4.2.11 Agreement in Kriyol: The modifier.....	121
4.2.12 The case of the Guinean community of Lisbon .....	122
4.2.13 The case of numerals.....	123
4.2.14 The quantifiers .....	124
<b>4.3 Bare noun phrases: Distribution and interpretation .....</b>	<b>126</b>
4.3.1 Subject bare nouns .....	127
4.3.2 Object bare nouns.....	131
4.3.3 Non-argument bare nouns .....	134
4.3.4 Bare nouns as predicates.....	136
4.3.5 The copula system of Guinea-Bissau Creole: An overview .....	139
<b>4.4 Derivation of BN interpretation.....</b>	<b>142</b>



4.4.1	How aspect influences BN interpretation.....	142
4.4.2	Predicate types and the interpretation of bare subjects .....	144
4.5	Summary and conclusion.....	145
<b>Chapter 5.....</b>		<b>147</b>
<b>Crosslinguistic comparison of bare nouns .....</b>		<b>147</b>
5.1	<b>Cape Verdean Creole .....</b>	<b>147</b>
5.1.1	The nominal and determiner system of CVC .....	147
5.1.2	Bare nouns in CVC.....	151
5.2	<b>The Creole of São Tomé .....</b>	<b>153</b>
5.2.1	The nominal system of Santome .....	154
5.2.2	Bare nouns in Santome .....	158
5.3	<b>Papiamentu.....</b>	<b>160</b>
5.3.1	The nominal system of Papiamentu .....	161
5.3.2	Bare nouns in Papiamentu.....	165
5.4	<b>Brazilian Portuguese.....</b>	<b>169</b>
5.4.1	The determiner system of Brazilian Portuguese .....	171
5.4.2	Bare nouns in Brazilian Portuguese .....	174
5.5	<b>Summary and comparison: BNs among creoles .....</b>	<b>179</b>
5.6	<b>Mandarin Chinese .....</b>	<b>182</b>
5.6.1	On word order in Mandarin .....	183
5.6.2	Chinese nominal system: Count vs. mass distinction .....	184
5.6.3	Bare Nouns in Mandarin .....	188
5.7	<b>Vietnamese .....</b>	<b>191</b>
5.7.1	The nominal system of Vietnamese .....	191

5.7.2	Vietnamese bare nouns .....	199
5.8	Gbe languages .....	202
5.8.1	The nominal system of Gbe languages .....	203
5.8.2	Bare nouns in Gungbe and Fongbe .....	210
5.9	Summary and comparison: BNs among non-creoles .....	213
Chapter 6.....		216
Semantic number: Theoretical implications on bare nouns.....		216
6.1	Depréz's Plural Parameter .....	216
6.2	The syntax of bare nominals .....	218
6.3	Derivation of BN interpretation within Depréz's account .....	219
6.4	Implications of the Plural Parameter: A comparison between +PL and -PL languages .....	221
6.4.1	+PL languages: The case of English, French and Italian .....	222
6.4.2	-PL languages: French-based and Portuguese-based creoles.....	224
6.4.2.1	Haitian Creole.....	224
6.4.2.2	Portuguese-based Creoles: CVC and GBC.....	226
6.5	An alternative approach to Depréz's proposal: Introductory remarks.....	228
6.5.1	An alternative account.....	228
6.6	An overview of the features pertaining to different types of languages with respect to the Num requirement.....	230
6.7	Predictions based on the above discussion relative to each group .....	234
6.8	Ways of forming NumPs .....	237

<b>6.9 Taxonomy .....</b>	<b>240</b>
<b>CHAPTER 7.....</b>	<b>244</b>
<b>Conclusions .....</b>	<b>244</b>
<b>7.1 Summary and generalizations.....</b>	<b>244</b>
<b>7.2 Remaining problems .....</b>	<b>250</b>
<b>References.....</b>	<b>251</b>

## List of abbreviations

1sg	first singular person
2sg	second singular person
3sg	third singular person
1pl	first plural person
2pl	second plural person
3pl	third plural person
AdjP	adjective phrase
AdvP	adverbial phrase
ANT	anteriority
Asp	aspect
BN	bare noun
BNP	bare noun phrase
BP	Brazilian Portuguese
CL	classifier
CONT	continuous
COP	copula
CVC	Cape Verdean Creole
DEF	definite
DEM	demonstrative
DP	determiner phrase
Eng	English
EP	European Portuguese
fem	female
FOC	focus
Fr	French
FUT	future
GBC	Guinea-Bissau Creole
Ger	German
HAB	habitual

IMPF	imperfective
INDEF	indefinite determiner
inf.	infinitive
INTERJ	interjection
It	Italian
L2	second language
masc	masculine
N	noun
NEG	negation
NP	noun phrase
NumP	number phrase
PAST	past marker
PERF	perfective
PL	plural
POSS	possessive
POST	posteriority
PP	prepositional phrase
PREP	preposition
PROG	progressive
Q	quantifier
SetP	set phrase
SG	singular
SPEC	specific
spec.fut.	specific future
TMA	tense-mood-aspect
TOP	topic

# Chapter 1

## Introduction

### 1.1 Goals of the dissertation

The present study has three main goals. First, it aims at being an exhaustive description of bare noun phrases in the Portuguese-based creole language spoken in Guinea-Bissau (West Africa): the perspective will be a synchronic one. Second, it will undertake a crosslinguistic comparison between a number of creole and noncreole languages in order to take into account crosslinguistic variation. The last aim is theoretical and typological in nature: this dissertation aims, in fact, to look at the semantics and syntax of Number crosslinguistically. We will create a taxonomy of languages on the basis of the different ways they form number phrases.

Bare noun phrases are nouns without any overt determiner. In the case of Guinea-Bissau Creole (GBC), or simply Kriyol, we will need a more restrictive definition: bare nouns in this language are, in fact, unspecified as for number and, as a consequence, may yield both singular and plural interpretations. In the present work, bare nouns will be defined as follows: nouns lacking any overt determiner and any number specification. This definition will hold also for both the creole and the noncreole group of languages involved in the comparison.

Both the distribution and interpretation of bare nouns (BNs) in Kriyol will be taken into account: BNs may, in fact, show up in both subject and object position. In principle, they may yield all possible readings: kind-referring, generic, existential, (in)definite and (non)specific. As a general tendency, subject BNs in Kriyol preferentially yield a definite reading. As for the direct object position, i.e. patient, no reading seems to be banned. Crucially, the plural specific reading is unlikely. On the other hand, bare indirect objects, i.e. recipient, yield a definite reading. As for nonargument positions, such as topicalized, left- and right-dislocated, and clefted BNs, they usually yield a definite reading; as for bare prepositional phrases (PPs), no reading seems to be banned. Finally, predicate BNs may yield both singular and plural interpretations.

Along with contextual cues, factors which play a role in the derivation of the correct interpretation of bare nouns in Kriyol are aspect and predicate type. As for aspect, bare objects in past perfective contexts yield definite (specific) readings; however, the indefinite reading is not excluded (the context must be suitable). Bare objects in imperfective (continuous) contexts with the continuous aspect marker *na* yield definite (specific) readings. Finally, bare objects in imperfective (habitual) contexts with the habitual aspect marker *ta* trigger indefinite nonspecific interpretations. On the other hand, as for predicate types, bare subjects of stage-level predicates yield existential readings, whereas bare subjects of individual-level predicates derive definite generic readings.

The description of Kriyol bare nouns will be inserted in a broader perspective: we will, in fact, look at the nominal and determiner system of Kriyol. The plural marker, the indefinite determiner and the demonstrative system are some of the issues that we will look at in detail. Both the description of bare nouns and the nominal and determiner system of Guinea-Bissau Creole will be dealt with in Chapter 4.

As for the crosslinguistic comparison, we will take into account both creole and noncreole languages. The first part of the comparison will include Cape Verdean Creole and Santome (Portuguese-based creoles spoken on the Cape Verde islands and on the island of São Tomé, respectively), Papiamentu (Spanish/Portuguese-based creole spoken on the ABC islands, i.e. Aruba, Bonaire and Curaçao), and Brazilian Portuguese. This latter is inserted in the comparison on the basis of the fact that, as a similarity to the creoles mentioned above, it has a wide distribution of bare nouns. Moreover, Brazilian Portuguese is often described as a semicreole (Holm 1992).

The second part of the comparison, on the other hand, will take into account typologically different languages: classifier languages such as Chinese Mandarin and Vietnamese, on the one hand, and Gbe languages such as Gungbe and Fongbe, on the other. Both the crosscreolistic and the crosslinguistic comparison will be dealt with in Chapter 5.

As for number interpretation, Kriyol bare nouns, and creoles in general, may yield both singular and plural readings; they may furthermore be unspecified as for number. This will turn out to be crucial for the theoretical discussion of Chapter 6 and, consequently, for the taxonomy of languages that will derive from the syntactic-semantic model developed in this work. This approach is inspired by Depréz's (2005, 2006, 2007) model, namely the so-called Plural Parameter.

## **1.2 Methodology**

The description of the nominal system of Kriyol is based partly on Kihm's work (1994, 2007) and partly on my corpus, its analysis and the consequent observations. The analysis of bare nouns is based on my corpus. It consists of six hours of spontaneous speech: they were recorded and transcribed. Only three of the six hours could be verified together with my informants; as a consequence, I decided to use only these three hours for the present thesis. The verified corpus consists of approximately 25.000 words; the transcriptions were translated into English and supplied with glosses. This material is the result of four interview sessions: two of them were recorded during my fieldwork among the Guinean community in Berlin (Germany), and the other two during my fieldwork among the Guinean community in Lisbon (Portugal). The informants in Berlin are 40 and 61 years old; they come from Buba (Southern part of Guinea-Bissau) and Bula (in the middle of the country, not far from Bissau), respectively. The two informants in Lisbon are 37 and 28 years old; the first one comes from Bissau, and the second one from Bissorã (in the East). I could not journey to Guinea-Bissau because of the unstable political situation.

## **1.3 Outline of the dissertation**

The present dissertation is divided into seven chapters, with the first one serving as an introduction to the organization of the work. Chapter 2 is a general introduction to the creole language of Guinea-Bissau. We will briefly review some historical facts relative to Portuguese colonialism in West Africa. Moreover, we will give an overview of the scenario responsible for the emergence of this creole; we will furthermore look at the different theories of creolization. Finally, we will give a general description of the morphosyntax of this creole.

Chapter 3 is an introduction to the study of bare nominals: both semantic and syntactic facts are taken into account. The semantic categories of Definiteness, Genericity, and Specificity will be dealt with: these are, in fact, crucial for the domain of nominal interpretation. As a consequence, they are usually lexicalized in natural languages. In Chapter 3 we will also look at some of the major semantic and syntactic approaches concerning bare nouns. As for the semantic models, we will present Carlson's (1977) treatment of English bare plurals, Partee's (1986) model of type-shifting operations, and Chierchia's (1998)



Nominal Mapping Parameter. As for the syntactic perspective, we will describe Longobardi's (1994) model of N-raising; on the basis of Longobardi's work, we will also give an overview of bare nouns in Romance and Germanic. Moreover, we will introduce Depréz's (2007) Plural Parameter, which will be the basis of our analysis in Chapter 6. Finally, we will briefly treat the case of bare nouns in European Portuguese, Kriyol's lexifier.

Chapter 4 aims at representing an exhaustive description of the nominal and determiner system of Guinea-Bissau Creole, based on the work of Kihm (1994, 2007) and on the analysis of data from my corpus. Particular emphasis will be given to BNs: we will look at both the distribution and interpretation of Kriyol bare nouns. It turns out that bare nouns in this language may yield a wide range of reading: kind-referring, generic, existential, (in)definite and (non)specific.

Chapter 5 is a crosslinguistic comparison of the distribution and interpretation of BNs in a selected number of languages. The study is based on data from the relevant literature. To better understand the behavior of BNs in the languages involved in the comparison, we will also describe their nominal and determiner systems. As we mentioned above, the first part of Chapter 5 will deal with Cape Verdean Creole, Santome, Papiamentu and Brazilian Portuguese. Although there are some important differences among these languages, they seem to behave in quite a homogeneous manner. On the other hand, the second part of the comparison will take into account Mandarin Chinese, Vietnamese and Gungbe and Fongbe (Gbe cluster of the Kwa group of the Niger-Congo language family). These noncreole languages show a free use of bare nouns, similar to the creole case. Similarities and differences are taken into account; all languages dealt with in this chapter can be said to share (almost) the same distribution and interpretation as for bare nouns.

In Chapter 6 we will discuss the theoretical consequences of the present crosslinguistic study of BNs. The theoretical implications discussed here concern the category Number. The chapter is divided into two main sections: the first describes the Plural Parameter as proposed by Depréz (2005, 2006, 2007), whereas the second section proposes an alternative model and a taxonomy of languages with respect to the (non)overt realization of the category Number. The model proposed in Chapter 6 is inspired by Depréz's Plural Parameter.

Finally, in Chapter 7 we summarize the findings of the present work with respect to GBC, the crosslinguistic comparison, and the theoretical implications about the category Number.

## Chapter 2

### A Sketch of Kriyol: Historical, social and linguistic aspects

#### 2.1 Lusofonia

During the last five centuries a sort of ‘Portuguese linguistic empire’ has arisen. It includes about 210 million speakers all over the world and extends from Sri Lanka to Japan, to Malaysia, to Brazil and to some West African countries, like Guinea-Bissau, the region of Casamance (in the southern part of Senegal) and Cape Verde islands. Nowadays, Portuguese is the official language of several independent countries: Portugal, Brazil, Angola, Mozambique, Guinea-Bissau, Equatorial Guinea, Cape Verde, São Tomé and Príncipe, Macao and East Timor. It is also a minority language in countries like Andorra, Luxemburg, Swiss and South Africa. This state of things is expressed by the Portuguese noun *lusofonia*. It refers to a space of cultural, linguistic and philosophical heritage of communities and countries that communicate in Portuguese.

In the 1420s, when the Portuguese set sail southward from the seaports of Lisbon, Sagres, and Lagos, they triggered a chain reaction which is responsible for the vast spread of the Portuguese language and for the emergence of Portuguese-based creoles. As Kihm (1994, 2011) reports, they had not planned for this to happen. On the contrary, during an era when the Moors ruled from the commercial point of view, they just wanted to create new markets and assault Islam from the rear. They sailed along the whole western coast of Africa. Then they proceeded southward through Sierra Leone and Zaire, until the Cape of Good Hope in today’s South Africa. Later they reached Mozambique on the east coast of Africa and from there they went to India. They did not limit their ventures to the south and east, but they also reached the territories of today’s Brazil. By the end of the 15<sup>th</sup> century the Portuguese had conquered not only a huge territorial dominion all around the world, but they had also set the basis for the modern geography. From then on Portuguese has been one of the most widely spoken languages in the world.

## 2.2 The rising of a new typology of languages: The creoles

Portuguese and other European languages like Spanish, French, English and Dutch had a crucial importance as a superstrate in the formation of a new species of languages: the creoles. About 20 Portuguese-based creoles have been attested around the world, seven of which are in West Africa: *Kriyol* of Guinea-Bissau, Casamance Creole (or Ziguinchor Creole),<sup>1</sup> *Criol* of Cape Verde Islands, *Angolar* and *Forro* in the island of São Tomé, *Annobonese* in the island of Annobon, and *Principense* in the island of Principe.

At the end of the 15<sup>th</sup> century, other Europeans followed the Portuguese footsteps and left bound for Africa, North and South America, and Asia. Spaniards, Englishmen, Frenchmen, Dutch, and Portuguese: they all have been responsible for the spread of their languages throughout the world and for their contact with those of the dominated populations. However, they are also responsible for the world's first modern genocide, committed in the New World's plantations and during the transportation of the slaves from their native countries.

It was exactly in scenarios like colonial plantations and hard labour, slavery and trade, that two new kinds of language were born: pidgins and creoles, both contact languages. Direct consequences of this are new contact situations, where different groups of speakers met each other for commercial purposes. Since the communication was limited to those purposes and occasions and was not an everyday linguistic exchange, no group of speakers had the opportunity and the occasion to learn the language of the other group(s). Furthermore, there was no phenomenon of bi- or multilingualism; therefore, there was the need for a common language, which could facilitate the interaction between the different groups of speakers. As a consequence, pidgins and creoles emerged. The greatest contribution to the vocabulary of those contact languages came from the language spoken by the socially dominant group: henceforth *superstrate* or *lexifier* language. On the other side, the *substrate* languages – so the native languages of the slaves are called – seem to have also played a role in the formation of the arising variety.<sup>2</sup>

---

<sup>1</sup> In Intumbo, Inverno & Holm (2013) Casamance creole is treated as a separate language.

<sup>2</sup> In every language contact situation, with *superstrate* or lat. *superstratum* one means the language spoken by a dominant group that has influenced another group from a population subordinate to it. On the other hand, *substrate* or lat. *substratum* indicates the language(s) spoken by some populations, which has influenced that of a group by which they were dominated. See Matthews (2007:390, 392).

Before we continue with our discussion of creole genesis and the various hypotheses on this, we should note that there are several types of pidgins and creoles. As for the first type of language, i.e. the pidgins, Mühlhäusler (1986) distinguishes among jargons, stable pidgins and expanded pidgins. As a crucial difference from stable and expanded pidgins, jargons are rather unstable. In contrast to creoles, pidgins do not have native speakers; however, expanded pidgins are getting native.<sup>3</sup> Now we come to the case of creoles. Similarly to pidgins, creoles may be distinguished on the basis of how they emerged: we have plantation, fort, and maroon creoles (Bickerton 1988). In fact, some creoles emerged in plantations (e.g. Caribbean creoles such as Jamaican and Haitian); some others emerged at the forts in the western coast of Africa and their surroundings (e.g. Kriyol and Kinubi). Finally, maroon creoles developed thanks to slaves who escaped plantations and formed isolated communities where their creoles emerged (e.g. Saramaccan, Palenquero and Angolar).<sup>4</sup>

With regard to the issue of superstrate and substrate languages we introduced above, we need to say that although they seem to have had a relatively limited influence in cases like that of Kriyol, some scholars maintain that the substrate languages have played such a crucial role that they have based a theory of Creole genesis on the substrate languages, as we will see below.

As far as it regards the pidgins<sup>5</sup>, their vocabulary is limited to the contact situations. Therefore, it is not extendable to the everyday communication. The same is true for the morphosyntactic resources: they lack systematic structures that can be traced back to a unique parent language. They would be rather an interlinguistic ‘compromise’ between the languages that took place in the exchange. In this sense Thomason (1997) claims that pidgins and creoles have a non-genetic development<sup>6</sup> because their structures cannot be ascribed to one language only. They would have been subject to a ‘language split’, which broke their relationship with the languages from which they were derived. Further evidence for a non-genetic development of pidgins lies in the fact that they have no native speaker: because of its limited functions, the pidgin has no native speaker and is the primary language of none of the groups involved in the exchange. Worth noting is also the dependency of the pidgin from the

---

<sup>3</sup> See Pieter Muysken and Norval Smith (1994) and Peter Bakker (1994) for more details.

<sup>4</sup> See Jacques Arends (1994) for a more detailed description.

<sup>5</sup> The term *pidgin* seems to derive either from the Chinese mispronunciation of the word *business* or from *pigeon*, a bird often used in the past for bringing written messages. See *Online Etymology Dictionary* and Bakker (1994:25).

<sup>6</sup> ‘Genetic linguistics’ is a standard notion in historical linguistics. Its core idea is that a language split brings to a new language family, where the new languages are changed forms of the languages from which they derive. See Thomason (1997:74).

social conditions: the pidgin changes according to the fluctuation in the social situation, whereas if the situation is stable in the time, a completely crystallized pidgin will develop. If other changes occur, then the crystallized pidgin will change too.

When (and if) a pidgin becomes more stable and plays the role of primary language in the contact community, it can develop into a creole. Indicative of such a change is the fact that the “new” language begins to have native speakers: by tradition this would be the most important factor distinguishing between pidgins and creoles. A further distinctive factor is the enhancement of the linguistic components of the new variety. The new social environment is no longer limited to a few communicative purposes; then the speakers have to face primary communicative functions that a pidgin could not face. Hence, the vocabulary has grown so that it can satisfy the new communicative situation. Another important factor here is the ‘review’ of the other linguistic components: morphosyntax, phonology and semantics. Once more we have to speak here of a linguistic compromise, where we rarely find marked characteristics that on the contrary we easily find in pidgins. Since creoles are in most cases the result of the contact between typologically very different languages (usually a European language, on the one side, and more non-European languages, on the other), the compromise mechanism, or radical re-shaping<sup>7</sup>, cancels the most marked characteristics of the languages involved in the phenomenon. That is why we are more likely to find creoles lacking inflectional morphology, compared to pidgins. Also Bickerton (1990) shares this idea, claiming that while pidgins are created by adults who learn a second language imperfectly, creoles would be on the contrary created by the children. They would use for that purpose an innate linguistic bio-program in order to supply the language with the necessary grammatical set.

A more recent approach to the question relating the birth of creoles and pidgins is based on the idea of demarcation on social factors, like communication environment, demography and other development circumstances (cf. Baker 1996, 2000, among others). Both creoles and pidgins would be created by adults, and not by children. The structural differences between the two contact varieties would lie in the different situations where the languages are used and in the distinct communicative purposes they have to fulfil. The nature of the inputs that take part in the ‘formation’ mechanism plays a crucial role. In fact, pidgins would develop in essentially commercial environments, where the speakers have limited exchange possibilities with each other. On the other hand, creoles would develop in

---

<sup>7</sup> According to Winford (1997:5-6), it is a sine-qua-non condition for the development of a creole/pidgin, which results from a radical reshaping of the structures of the languages involved in the contact.

environments like plantations where the speakers are in close contact with each other, but not with the colonizers. For Mufwene (1991), contact languages have to be classified on the basis of their development conditions, and not on the basis of structural factors. Following this approach, the terms ‘creole’ and ‘pidgin’ would have a mere historical significance; we should just talk of ‘contact languages’ that are distinguished from each other on the basis of their degree of restructuring<sup>8</sup>. In the opinion of Myers-Scotton (2002), the development of creoles has a lot of commonalities with the bilingual code-switching<sup>9</sup>: creoles would have the same mixture of inputs from the matrix language and the embedded languages.

### 2.2.1 Theories of creolization

The origin and development of creole languages are still debated, and therefore the controversial topic is also extremely fascinating. There are many theories about their origins: some of them are based on the European input, others on non-European inputs. Some have as their core question the nativization<sup>10</sup> of a pidgin, from which a creole would develop, whereas more recent ones focalize on universal principles.

Between the 1960s and 1970s the monogenetic hypothesis, first elaborated by Taylor (1961) and Thompson (1961), was the most supported. There are two versions of it: following the first one, all Creole languages would have developed from a Portuguese-based pidgin spoken between the 15<sup>th</sup> and 18<sup>th</sup> centuries in the commercial colonies and forts of West Africa. Also creoles with a lexifier language other than Portuguese would have developed from this unique pidgin via relexification: this could explain the many commonalities among creoles. The second version of the monogenetic theory digs more deeply and claims that the pidgin, from which the creoles – at least the Atlantic ones – would have developed, derives from the *lingua franca*<sup>11</sup> spoken in the Mediterranean since the Crusades until the end of the

---

<sup>8</sup> The term *restructuring* describes a syntactic or other process by which the structure assigned to a form is changed without change to the form itself. See Matthews (2007:345).

<sup>9</sup> *Code switching* refers to the case of bi- or multilingual speakers that use - syntactically and phonologically correct - elements of the two or more varieties/languages during a conversation without rising confusion or mixture. The latter case is known as *code mixing*.

<sup>10</sup> The term *nativization* refers to the process by which a language has become *nativized* in a community. This means that the language under question is now the native language of the group of speakers. See Matthews (2007:258).

<sup>11</sup> The Mediterranean *lingua franca* was called *sabir*, mispronunciation of the Catalan verb *saber*, ‘to know’. There was several diatopic and diachronic variants, but the most widespread consisted of a mixture between Italian (for the most part, the Italian varieties of the Maritime Republics of Venice and Genoa)

19<sup>th</sup> century. The monogenetic theory initially had great success, but then it began to be considered unreliable and was replaced by a ‘reduced’ monogenesis: English and French creoles would derive respectively from an English pidgin and a French one, both coming from West Africa.

Following Holm (1988), the hypotheses of scholars like Naro, Schuchardt, Seuren and Wekker consist of a completely different point of view: they hold that creoles would have evolved from a ‘foreigner talk’ variety, which the native speakers of the lexifier language spoke with the indigenes<sup>12</sup>; the Creole would then be shaped on this variety. However, this hypothesis does not seem to be reliable: on the contrary, it rather seems to be more probable that the linguistic varieties learned by the foreigners reflect the particular structures of the creole languages. As an alternative to this theory, the 19<sup>th</sup> century Portuguese linguist Adolfo Coelho proposes the imperfect learning of L2. Research in second language acquisition have discovered several features of ‘interlingual systems’ that can be found in pidgins and creoles as well: unvarying verbal forms, lack of determiners – or use of demonstratives instead of determiners -, negation always in preverbal position, use of adverbs in order to express modality, a fixed word order, and finally, no – or at least reduced – plural marker. Summarizing, pidgins and creoles would result from an imperfect L2 learning of the lexifier language by the slaves.

As far as it concerns the non-European input in the Creole genesis, the most important hypothesis is based on the linguistic substrate and on ‘relexification’: the substrate language(s) of the slaves would have somehow influenced the new emerging variety, i.e. the creole. Lefebvre (1998) integrates the substratist approach with the relexification hypothesis, claiming that the substrate languages together with a European language would have been relexified causing the rise of the new variety.

Among the ‘universalist’ hypotheses, we find the theory of Coelho: the Romance-based creoles as well as the Indo-Portuguese dialects<sup>13</sup> would have been developed thanks to physiological or psychological laws that act the same way everywhere. Hence, for Coelho the

---

and Spanish with Sicilian, Greek, Turkish and Arabic influences. The expression *lingua franca* derives from the Arabic *lisan-al-faranġi*, lit. ‘European language’.

<sup>12</sup> Crucial agents in the formation of a ‘foreigner talk’, of a creole based on it, of a ‘baby talk’ and more generally in the formation of a simplified language would be linguistic universals that a community of speakers uses when it needs simplified registers in order to communicate with individuals that cannot properly understand the language: foreigners, babies etc. See C.A. Ferguson (quoted in J. Holm 1988:62).

<sup>13</sup> The Indo-Portuguese dialects originated from the Portuguese control of parts of the coast of India and of Ceylon from the beginning of the 16<sup>th</sup> until the mid-17<sup>th</sup> century. These dialects are divided into two groups: ‘Gauro-Portuguese’ – influenced by Indic languages - and ‘Dravido-Portuguese’ – influenced by Dravidic languages. See Holm (1988:284-285).

origin of Creole languages would not depend on the languages taking part in the contact situation. A crucial role in the universalist hypothesis is played by the 1970s 'Language Bioprogram Hypothesis' of Bickerton, although its influence on creole studies is rather indirect: it has, in fact, generated a lot of studies and researches. Scenarios of creole genesis would be the plantations, where lots of slaves coming from different ethnolinguistic communities were deported and forced to work. Here and under the conditions described above, the pidginization of the variety spoken by the dominant group – the European one – took place: the pidgin would be, thus, the result of this promiscuous situation, where every ethnic group had its native language, different from the other, and both communication and mutual intelligibility would have been impossible without the pidgin. Following Bickerton, the children born in such a situation would have realized, on the one hand, that they could not use the language of their parents in order to face their need for communication but, on the other hand, they could not use the pidgin as well. Bickerton compares this situation to that of the primitives that would have had the necessary cognitive structures in order to create a base language: at first they would have lexicalized general simple concepts and then would have reached the distinction subject/object. From then on, a stable subject-verb-object word order and the Tense-Mood-Aspect system of the verb would have been achieved. In Bickerton's words:

*This brings us to the state of full-fledged, albeit basic, language – one which reaches the complexity of most creoles. It is here, then, that the 'biological development' of language ceased and the cultural development began.*

(Black and Gilbert 1991:112)

As far as it concerns creoles, Bickerton claims that it is the cultural overlap that distinguishes the creoles from the other languages, whereas the addition of culture spurs the decreolization<sup>14</sup> of creoles. Then the children of the plantations would have acquired the vocabulary from the pidginized varieties and, in order to create a proper language, they would have drawn from their linguistic bioprogram or innate linguistic capacity.<sup>15</sup> The language

---

<sup>14</sup> By *decreolization* is meant the historical process by which a creole is progressively assimilated to a standard language. See Matthews (2007:83). In the case of Kriyol, it means that if the continuum with Portuguese continues to increase, this creole could be assimilated to Portuguese. See also note 28 below. This risk seems to be far away, as for Kriyol (Kihm 1994).

<sup>15</sup> Bickerton's *Paradox of Continuity* says that the language should have developed from a preexisting system, although such a system does not seem to be traceable. In order to overcome this paradox is to consider the language as a representational system, hence a mechanism that created in most part its own inputs by itself. The language would be then not only able to increase the things human beings could talk of, but also to increase the types of these things. No other communicative mechanism could have done it. See D. Bickerton (1990:8, 75).



described so far is called ‘radical creole’, and its grammar would be near to the non-marked state of our innate *faculté de langage*.<sup>16</sup> Subsequent studies, however, have shown that creoles are not as similar to each other as Bickerton thought, that creolization has taken more time – that is why a lot of creolists nowadays agree with the ‘gradual basilectalization’<sup>17</sup> theory – and that some creoles have been influenced by their substrate languages.

Before concluding the discourse on the universalist approaches, another issue should be taken into account: the kinds of universals. Following Muysken and Veenstra (2005), the linguistic universals that take part in the formation of a pidgin/creole are of two kinds: procedural and constitutive. The first kind consists of the set of strategies used by the speaker in contact situations, like L2 learning or grammaticalization<sup>18</sup>, reduction<sup>19</sup> and ‘semantic transparency’. This latter one has been initially described by Naro as follows: “express every meaning element, perceived as separated from the others, with a separated stressed form” (Muysken and Veenstra 1995:122). Later, Seuren and Wekker made of Naro’s principle a ‘theory of semantic transparency’: the semantic structures of creoles reflect the universal semantic structures and are directly mapped onto the surface structure without complicated intermediary relations. The second type of universals is the constitutive one and consists in the universal properties of pidgins/creoles that result from the formation of such languages: preverbal particles like negation and aspect markers, the rigid SVO word order, serial verbs, transparent systems of interrogative words, morphologically complex reflexives, plural marker involving third person plural pronoun, generalized locative preposition, fronting rules – focus and predicate split – and presence of double accusative (ibid.). These kinds of universals belong to the generativist research field and go along with the claim that each natural language has to conform to the Universal Grammar, whose non-marked option would be represented by the creoles.

---

<sup>16</sup> In the Generativist Theory, *faculté de langage* is the capacity – specific of the human kind – which enables the acquisition of the language. It derives from the mentalist conception developed by philosophers like Descartes and Humboldt.

<sup>17</sup> As Baker (2000) reports, this hypothesis was first conceived by Chaudenson (1992) in relation to French-based creoles and was then applied by Mufwene (1996) to other creoles. Following this hypothesis, the first plantation slaves would have acquired a sort of approximated variety of the dominant group’s language, that of the Europeans. Every new wave of slaves’ deportations into these scenarios brought to a further approximation in the learning of the target language. The result of such a process would have been the development of a basilect-creole.

<sup>18</sup> *Grammaticalization* is the process by which a unit with lexical meaning – any aspect of meaning that is explained as part of a lexical entry for an individual unit – changes into one with a grammatical meaning – any aspect of meaning described as part of the syntax and morphology of a language as distinct from its lexicon. See Matthews (2007:164;224).

<sup>19</sup> An example of reduction is the fact that the Romance pre-verbal clitic had disappeared in creoles: *Èl me mira* (Spanish) vs. *E ta mira mi* (Papiamentu). See Muysken and Veenstra (1995:122).

## 2.3 The emergence of the Creole language of Guinea-Bissau

Before we lay out a historical overview of Kriyol, we should take note of an important fact: the creole of Guinea-Bissau and the one of Senegalese Casamance have often been described as two varieties of the same language. In the present work, we are going to look at the case of Guinea-Bissau Creole only. GBC is a fort creole, after the subdivision of creoles described above in 2.2 (Kihm 2011).

Between 1445 and 1446, after having colonized Madeira and the Canary Islands and the Archipelago of the Azores, the Portuguese reached Senegal and the Cape Verde Islands, the coast of Gambia, Casamance and Guinea-Bissau. Since the mid-15<sup>th</sup> century the Portuguese mainly used these colonies to enhance their own commerce and as a source of slaves: native Africans from these regions were deported to Portugal and employed as domestic slaves. The slaves were “rather free to move about and to mix with the white population into which they merged without leaving any recognizable phenotypical trace” (Kihm 1994:3).

As for the origin of Kriyol, the matter is quite intricate and there are several opinions. We will briefly introduce them and then present more accurately Kihm’s (2011) treatment of this issue. According to Rougé (1986), a ‘proto-Kriyol’ would have existed by the end of the 15<sup>th</sup> century. As far as the variety of Portuguese spoken by the deported slaves, Naro (1978) claims that they spoke an ‘artfully devised pidgin’ – in Naro’s words, ‘a reconnaissance language’ (Kihm 1994:3). The Portuguese would have used this pidgin for the necessary communication with the slaves and to use some of them as interpreters (*linguas*) in subsequent expeditions. However, Goodman (1987) and Clements (1992) claim that the slaves in Portugal talked to each other in a foreigner-talk variety of Portuguese. According to Kihm (1994, 2011) such a *língua de preto* (Portuguese expression meaning ‘Black Portuguese’; also called ‘Portuguese Pidgin’) would have been the Basic Variety<sup>20</sup> of Portuguese spoken by the African interpreters and could have been the basis for the creoles of Senegambia<sup>21</sup> and Cape Verde. Evidence of that would be the language used by the actors playing the role of the slaves in 16<sup>th</sup>-century Portuguese comedies. This theatre language and the Portuguese-based creoles of West Africa share important similarities, although limited;

---

<sup>20</sup> These African interpreters have stayed no more than a decade in Portugal. Therefore, they did not learn properly Portuguese, but a basic variety of it (Kihm 2011:85).

<sup>21</sup> The term Senegambia could create a certain degree of ambiguity. Here it stands for today’s territories of Senegal and Gambia. On the other hand, Senegambia was the name of a Confederation, created in 1982 with the aim of promoting cooperation between Senegal and Gambia. The Confederation was broken up from Senegal because Gambia had refused to integrate with Senegal.

they consist in phonological and morphosyntactic features that no Portuguese native speaker could have invented (Kihm 1994:4). On the basis of the similarity between Cape Verdean Creole and Guinea-Bissau Creole, Jacobs (2009) assumes that a common origin for CVC and GBC is possible: it is not clear yet whether the ‘proto-creole’ developed on the continent or on the Santiago island of Cape Verde.<sup>22</sup>

As we introduced above, Kihm (2011) assumes that Kriyol emerged from a basic variety of European Portuguese; in Kriyol’s emergence, substrate languages had a limited role. Following Kihm, when the Portuguese reached the present-day Guinea-Bissau, they met more or less the same peoples living there nowadays: Manjakus, Mankanyas, Papels, Balantas, and Diolas, among others. The latter group, i.e. the Diolas, live nowadays for the most part in Senegal. Each ethnic group has its own language: most of them are Atlantic languages, although there are also Mande languages like Mandinka;<sup>23</sup> all of them belong to the Niger-Congo phylum (2011:84). Nonetheless, the complex linguistic situation has never been a problem (Rougé 1986): “extended bilingualism was probably the solution, plus possibly the use of Mandinka as a trade language” (Kihm 2011:84). This means that local people did not need any language for communication purposes among different peoples. Portuguese presence was not strong in Guinea-Bissau: they never took real control of the land and furthermore and they did not reach into the interior. In fact, they used their colony for commercial purposes and slave trade. Slaves had often to act as interpreters: “these African interpreters [...], having stayed not more than a decade in Portugal, must have spoken, not ‘proper’ EP, but the Basic Variety of it that came to be known as *Língua de Preto* ‘language of the Black’” (Kihm 2011:85, emphasis in original). Thanks to their tasks, such as that of interpreters, slaves could earn their freedom: this could explain the “entrance for pidgin EP on the West African coast” (ibid.). The *lançados*<sup>24</sup> (lit. ‘rejected’) played a crucial role in the creolization: as Kihm (1994, 2011) explains, they were Portuguese outcasts, socially excluded because of their criminal record or Jewish parentage. They were smugglers who landed clandestinely on the coast and traded in slaves and other goods. They were easily accepted by the local population and mixed with it: their sons were called *filhos da terra* (lit. ‘children of the earth’ and also *mestizos* or *mulattos* for indicating their mixed origin), and over time they

---

<sup>22</sup> Jacobs assumes that the proto-creole from which both Cape Verdean and Kriyol have emerged, has developed on the Santiago Island. It would have been brought onto the continent on the occasion of the foundation of Cacheu (Jacobs 2009).

<sup>23</sup> Mandinka, and its various dialects, is spoken by the Mandings. These people ruled and conquered those areas of West Africa at the time when the Portuguese arrived.

<sup>24</sup> The term *lançados* ‘castaways’ comes from Portuguese *lançaram-se* and literally means “(who) threw themselves’ into the heart of darkness in order to trade with whoever they pleased” (Kihm 1994:4). It was, of course, illegal since the commerce was entirely in the hands of the Portuguese Crown.

formed “the bulk of the Portuguese presence in Guinea” (85).<sup>25</sup> A very important role in the emergence of Kriyol was played by the *grumetes* ‘shipboys’: they were Africans that had converted to Christianity and lived in *povações* ‘settlements’ or *aldeias* ‘villages’, special villages near the Portuguese outposts, like Cacheu or Ziguinchor<sup>26</sup>. They acted as mediators between the Portuguese and the other Africans.<sup>27</sup> They had thus an in-between position: on the one hand, they were able to communicate with the Portuguese group in the linguistic variety that was closest to EP, namely the emerging creole, and on the other hand, they could speak to the local people since they knew several local languages: it was an “emerging creole [...] rooted in the pidgin EP the *lançados* must have used with the villagers among whom they had settled” (86; emphasis in original). Following Kihm, Kriyol had the function of group identity: on the one hand, the *grumetes* never switched to EP probably in order to keep “the [...] Portuguese community at a profitable arm’s length”; the fact that they worked as intermediators between the Portuguese and the local populations let them space enough for their own returns. On the other hand, thanks to Kriyol, they were distinct from the other local peoples who identified “them with the economically powerful Portuguese community” (86). Concluding, Kihm assumes that “creolisation was actually a matter of restricting access in order to construct and preserve a separate community” (87).

### 2.3.1 Kriyol in more recent times

The first document written in Kriyol is from the 19<sup>th</sup> century; at that time there were more clear-cut differences among geographical dialects than today. Gomes and Mendonça (1981) and Pinto-Bull (1989) speak about the dialects of Cacheu, Geba and Bissau. However, in today’s Kriyol the only distinct geographical variety is that of Ziguinchor, outside the borders of Guinea-Bissau.<sup>28</sup>

---

<sup>25</sup> Until the beginning of the 20th century, very few Portuguese were present in the area corresponding to present-day Guinea-Bissau (Kihm 2011:85).

<sup>26</sup> Cacheu is today one of the most important cities in the North of Guinea-Bissau, whereas Ziguinchor is the principal city of Casamance, the Southern region of Senegal where Casamance creole is spoken. Cacheu was the first fortified settlement (*praça*) and was not founded until 1588 (Kihm 2011:84).

<sup>27</sup> They were called ‘shipboys’ because they often worked for the Portuguese as sailors or as mediators for trade relations (Kihm 2011:86).

<sup>28</sup> The Berlin Conference in 1884-1885 devolved the region of Casamance to the French government, so that this region became part of the French colony of Senegal. As a consequence, Kriyol here was influenced by Muslim Wolof-speakers, became the language of the Christian community in Senegal and were isolated from Kriyol of Guinea-Bissau. Some archaisms present in Casamance Creole, but completely disappeared in Guinea-Bissau Kriyol, would show it (Kihm 2011:87).

Between the end of the 19<sup>th</sup> and the beginning of the 20<sup>th</sup> century, Kriyol had a great expansion: the Portuguese had indeed decided to take real possession of their African colonies that they had used until then only as a source for slaves. It was the time of the Scramble for Africa, and the Portuguese worried about losing the African colonies that the Berlin Conference (1884-1885)<sup>29</sup> had officially recognized as territories of Portugal. Therefore, with the help of the *grumetes*, they implemented a series of military ‘pacification’ campaigns; four of them “were led from Major Teixeira Pinto and supported by a number of Fula chiefs and about 400 Fula and Mandinga troops, along with a number of foreign mercenaries between 1913 and 1915, which resulted in the destruction of a number of Balanta, Budjugu, Felup, Mandjaco, Oinca and Papel villages. The campaign clearly showed the growing collaboration between the Fula and the Portuguese in the rural areas of Portuguese Guinea, but the level of brutality can be gauged by the public outcry it evoked in Lisbon” (Lobban and Forrest 1988:107, in Vigh 2006). After these campaigns, the Portuguese gained entire control of the country<sup>30</sup>. Now the need for a lingua franca became urgent, but only Kriyol could play that role since none of the autochthonous languages had interethnic currency and Portuguese was not diffused among the populations<sup>31</sup>. So, Kriyol broke its *continuum*<sup>32</sup> with Portuguese and became a fully autonomous system, although Portuguese sometimes returned during the 20<sup>th</sup> century as a consequence of Catholic missions<sup>33</sup>.

In 1946 a Bill of Law decreed the criteria in order to distinguish among the population of the country on the basis of social factors. Whether they were indigenous or Portuguese citizens, it had to be established on the basis of a linguistic condition: whoever wanted to be

---

<sup>29</sup> The Berlin Conference of 1884-1885 was organized by Otto von Bismarck, first Chancellor of Germany, with the aim of regulating the European colonization and the trade in Africa.

<sup>30</sup> The military campaigns took place only in the Western part of the country, where “animist” populations lived (Manjakus, Papels, Balantas, etc.). In the Eastern part the Muslim Mandingos and Fulanis were long organized in powerful kingdoms and the Portuguese had to negotiate with them (Kihm 1994:269).

<sup>31</sup> In fact, the number of Portuguese in Guinea-Bissau was not sufficient in order to promote the use of Portuguese (Kihm 1994:6).

<sup>32</sup> The term *continuum* refers to a chain of linguistic varieties, whose phonology, morphosyntax and lexicon are strictly connected to each other. On the one end of this chain one finds the *basilect* and on the other end the *acrolect*. The former one is the lowest variety in terms of sociolinguistic prestige, whereas the latter one is the highest. The intermediary variety in this chain is called *mesolect*. A creole continuum may rise in the following situation: when a creole coexists with the lexifier language and when the creole speakers have a social reason for learning the standard variety (the lexifier language), acquiring then features of this variety to the detriment of the creole (Holm 1988:52). In the case of Kriyol, the continuum is a concatenation of varieties that connect *crioulo puro* (the oldest variety, as the speakers call it) to Portuguese.

<sup>33</sup> In 1940 Catholic missionaries came to Guinea-Bissau with the aim of supply primary education to the ‘indigenes’. The teaching language was Portuguese, and not Kriyol.

recognized as a Portuguese citizen had to be able to speak, read and write Portuguese<sup>34</sup>. From now on Kriyol had no more the same scholarly and literary recognition as before, although its use did not stop spreading.

Between 1961 and 1974 Kriyol had a new growth. It was the end of the colonial period and the beginning of the independence war, at the end of which Guinea-Bissau was internationally recognized as an independent country. During the war, Kriyol played in the PAIGC<sup>35</sup> the double role of tool for linguistic unification and symbol of a nationality: the Bissau-Guinean one. Following the plans, Kriyol had to become the language of the education and Portuguese had to remain the official language in order to keep the country in contact with the rest of the lusophone world. Although the plans did not materialize, Kriyol became the language for the every-day life in the capital city Bissau and the other principal cities, whereas each ethnic group speaks its own autochthonous language. In any case, Kriyol became the lingua franca for the whole country. Since gaining independence, Kriyol's literary production has increased: poetry, songs, and comic books. In the last years, two movies have been produced: Flora Gomez's *Nha fala* (Kriyol; lit. 'My voice'), 2002, and João Viana's *A batalha de Tabatô* (Port.; lit. "The battle of Tabatô"), 2013. The second film is actually in Mandinka and some parts are in Kriyol. Radio programs are normally in Kriyol, but TV programs are in Portuguese, with a massive presence of Brazilian *telenovelas*. Music production in Kriyol is rapidly increasing: not only traditional Guinean music, but also modern music such as rap and pop use Kriyol as their expressive means. Also, literature in Kriyol seems to be growing: several authors have started to write in Kriyol or, at least to put it close to Portuguese in their compositions; among young Guinean poets we find Edson Incopté with his poetry collection, *Insana Rebeldia*, and Michél Té, published in *V Antologia Lusofona de poetas losófonos*.

---

<sup>34</sup> There were also some economical conditions and a moral one: they had to demonstrate "good behaviour", in other words they had not to practice the "customs of the common people of their race" (Pinto-Bull 1989:108ff.).

<sup>35</sup> PAIGC stands for *Partido africano da independência de Guiné e Cabo Verde* (African Party for the independence of Guinea and Cape Verde). On 24<sup>th</sup> September 1973 the PAIGC declared the independence of Guinea-Bissau from Portugal. The independence of Guinea-Bissau was recognized only on 10<sup>th</sup> September of the following year, and until that moment, the country had been called 'Portuguese Guinea'.

### 2.3.2 Kriyol's *substratum*

The term 'substrate' (or lat. *substratum*) describes the language (or languages) spoken by the dominated population in a situation of language contact. As we saw above, according to one hypothesis of creolization or the other, the substrate is said to play a more or less important role in the formation of the new variety. In the present section, we are not going to take a position in the debate about creole genesis; nonetheless, we are going to look at some of the substrate languages of Guinea-Bissau Creole. The following description is based on the work of Kihm (2011).

Kriyol has a large number of languages as its substrate; among them, there are both Atlantic and Mande languages, whose grammatical features are very different from those of Kwa languages that are said by the substratists to have played a crucial role in the development of Atlantic creoles. However, we have to notice that African languages spoken in the areas of Guinea-Bissau and Senegalese Casamance do not seem to have had a crucial influence on the development of Kriyol grammatical features.

Kihm analyzes the grammatical features of Kriyol nominal and verbal domains that differ from those of the lexifier language, and compares them with those of a selected group of local Atlantic languages: Balanta, Diola, Mandinka, Manjaku and Wolof. With the exception of Mandinka, which is a Mande language, the others belong to the Atlantic cluster, are agglutinative and show a great morphological complexity, as to both derivation and inflection<sup>36</sup>. These features are certainly not shared by Kriyol.

As far as it concerns the noun phrase, we have first to deal with the category Number. Here Kriyol differs deeply from its substrate languages: Kriyol noun phrase is most of times bare and can be ambiguous between a singular and a plural interpretation. There is a way to overtly mark plurality in Kriyol, i.e. by adding to the noun phrase the suffix *-s* (or *-is/-es* if the word ends in a consonant). This kind of suffixation comes from European Portuguese, where the pluralization process is quite different from Kriyol: in Portuguese the mark of plurality is required whenever the noun phrase refers to more than one entity, whereas in Kriyol the plural marker appears on the noun only under conditions such as 'animacy' and 'referentiality'.<sup>37</sup> The situation in the substrate languages under investigation seems to be quite

---

<sup>36</sup> Mandinka shows a certain degree of complexity in derivational morphology (Kihm 2011:88).

<sup>37</sup> The first plurality condition, that of animacy, requires that the noun phrase has the feature [+human], whereas the condition of referentiality wants the noun to refer to a certain entity in order to be pluralized. We will describe these facts in more detail in Chapter 4 of the present work.

different from Kriyol: in fact, they all have an affix specialized for plurality such that they are required in each plural context, i.e. when more than one entity is referred to. In particular, Balanta, Diola, Manjaku and Wolof are noun class languages, “meaning that roots are realized as nouns via association with affixes, the two cumulative exponents of two morphosemantic features: Class and Number” (Kihm 2011:89-90). Here bare nouns can only be singular. Although it is not a noun class language, also Mandinka requires the pluralization of the noun phrase whenever more than one entity is referred to.

Another feature to take into account is the definite determiner. Kriyol has none, distinct from its lexifier language,<sup>38</sup> while Balanta and Manjaku pattern on a par with Kriyol in this respect, Diola and Wolof, on the contrary, have an overt determiner in order to distinguish between determinacy and non-determinacy. Also Mandinka has a way of expressing determinate reference: a specific determiner *-oo* which can be interpreted as definite or indefinite.

Among the substrate languages under study, we may find bare nouns in their strict definition as “nouns lacking both determiner and any specification for number” (Kihm 2011:145), only in Mandinka and Wolof: unlike in Kriyol, bare nouns in these languages appear in generic contexts only.. The kinds of interpretation ambiguity we may find in Kriyol, like [ $\pm$ ] definiteness or [ $\pm$ ] plural, cannot be found in any other local language.

As regards the indefinite determiner, substrate languages seem to have played here a certain role: in Kriyol the indefinite determiner is *un*, derived from the Portuguese indefinite determiner *um(a)*.<sup>39</sup> Both in Kriyol and in the local Atlantic languages there is the contrast between two kinds of indefiniteness: “the bare truly indefinite or generic nouns (‘any x’) and noun phrases involving an item that conveys specific indefiniteness (‘a certain x’)” (Kihm 2011:90). This contrast is fused in Kriyol into the indefinite determiner *un* with the particularity, in opposition to the other creoles, that the interpretation ‘a certain goat’ may also be conveyed by the expression *utru kabra*, lit. ‘(an)other goat’: the overlap of the meanings ‘other’ and ‘certain’ is typical of languages like Balanta and Manjaku.

---

<sup>38</sup> European Portuguese has specialized forms of definite determiner. Each of them is indeed specialized for gender and number: *o* and *a*, respectively for masculine and feminine singular, and their plural counterparts *os* and *as*.

<sup>39</sup> Exactly as for the case of the definite determiner, Portuguese offers a specialized form of the indefinite determiner for each combination between number and gender: *um/uma* ‘a, an’ (singular, masculine and feminine, respectively) vs. *uns/umas* ‘some’ (plural, masculine and feminine, respectively).



Before concluding the present section on the nominal domain, we have to spend some words on another typical grammatical feature of Kriyol, i.e. the absence of gender contrast. It is one more point where Kriyol differs from its lexifier European Portuguese: the latter one has indeed gender categories. Furthermore, once again Kriyol differs from its substrate languages insofar they are noun class languages. Nonetheless, the substrate could have had a certain influence in the fact that Kriyol did not take the binary gender contrast from Portuguese: the speakers of *lingua de preto* would have been competent in Atlantic noun class systems, being at least bilingual just as well as today, and this would have not helped them in acquiring the Portuguese gender system.

As we said above, also the verbal domain shows interesting phenomena, like the Tense-Mood-Aspect system. Kriyol and creoles in general have indeed a “special” TMA marking system, which is innovative in comparison with that of the lexifier language. At this point, we have to notice that GBC is different in this respect from the other Atlantic creoles, since it shows a different tense-aspect-mood system, namely aspect occurs preverbally, whereas tense and mood postverbally (there also cases of temporal markers which show up preverbally, but always at the right of aspect markers). Once again, it seems that substrate languages have played just a causal role in the development of the TMA system, and not a crucial one. As far as Kriyol aspect markers, as we will see, they are preverbal particles, *na* and *ta*, that bring an imperfective value. The aspect marker *na*, which brings a continuous interpretation to the event described by the verb, derives from Kriyol locative *na* < EP locative *na* (*in* ‘in’+ *a* ‘the.SG.FEM’) with the meaning of ‘on, in’. Kriyol form *na* + V derives from Romance construction “imperfective-locative copula + present participle of V / preposition *a* + V in its non-finite form”, and more specific from EP *estar fazendo/estar a fazer*, “to be (a-) doing”. The modern form *na* + V in Kriyol is a reduction of the form *sta na V* attested in old Kriyol texts and still in the Casamance variety of Kriyol (Kihm 2011:92-93). The substrate shows similar constructions as well, first of all in Balanta:

- 1) N ga a til-a.  
 1sg be in write-ing  
 ‘I am writing.’ (Intumbo 2007:65)

The sentence in (1) shows the same construction as in Kriyol: here *na* is replaced by Balanta preposition *a* and introduced by the verb *ga* ‘be’.

Furthermore, in Balanta as well as in Kriyol, the progressive construction of the verb may refer to a future event (Kihm 2011:93). This does not happen either in the other Atlantic languages or in EP. This could mean that Balanta has played a crucial role in the semantic evolution of *na*. Balanta could have had a certain influence, together with Manjaku, in the fact that the unmarked form of the Kriyol verb refers to perfective events or present states. We have now to specify that this feature is found in many other creoles; this does not seem to testify in favor of the substrate: “it could be that perfective or accomplished is the default Tense-Aspect value for event-denoting items, while imperfective is the default value for state-denoting ones” (Kihm 2011:94).

A point where Kriyol is not as typical as the other Romance-based creoles is the marking of anterior tense. Here Kriyol does not have a preverbal tense marker, but a postverbal one: *ba*. This morpheme also appears in Sotavento variety of Cape Verdean, but it is here a bound morpheme attached to the verb: it could be thus analyzed as an inflectional morpheme. In Kriyol, on the contrary, it is a free morpheme, which appears not only after verb phrases, but also in temporal locutions like *na kil tempu ba* ‘at that time’. The question of the origin of *ba* must be faced, if one wants to understand the actual state of things of this tense marker. It seems to derive from Portuguese *-va*, bound inflectional morpheme of first and third singular imperfect indicative. In Kriyol, *ba* has become a free morpheme, but maintained its location after the verb. Such a change to free form could have been driven by the substrate: nearly all substrate languages show a free morpheme conveying the same anteriority meaning as Kriyol *ba*. In Mandinka and Manjaku, the anteriority free forms are similar to that of Kriyol not only in the fact that they follow the verb, but also phonologically: indeed, the anteriority forms in Mandinka and Manjaku are respectively *ban* and *ba*. We have now to add that these anteriority markers are verbs with the meaning of ‘finish’ that are always postponed to the main verb they modify, as the following sentence from Manjaku will show:

- 2)        A-reala ba.  
             3sg eat finish  
             ‘S/He ate already.’ (Buis 1990:51)

In the example in (2), *ba* is postponed to the verb and is interpreted like a temporal adverb conveying the feature [+ anterior]. One may note here that the semantics is not the same as for Kriyol *ba*, which bears an imperfective meaning, but Manjaku shares with Kriyol the fact that the verb has been reanalyzed as an adverbial particle. Coming back to the question of the

origin of Kriyol *ba*, we have to add another possibility: it could also derive from the verb *kaba*, “to finish” < EP *acabar*. The path from *kaba* to *ba* could also be explained by the fact that Kriyol *kaba* was identified with Mandinka *ka ban* through lexical conflation (Kihm 2011:97), as attested by the form *kaba* in the following example from Kriyol:

- 3) Bajudas fididu jinjirba tudu i kaba, e say.  
 girls split gum all 3sg finish 3pl leave  
 ‘All the girls had their gums tattooed and then they left.’ (Kihm 1994:245)

The origin of *ba* seems thus to be still uncertain and swings between the Portuguese and the substrate explanation. There seems to be, however, more evidence in favor of the Portuguese origin of the anterior marker: against any “relexification hypothesis”<sup>40</sup>, it would thus derive from the lexifier, both in form and meaning, and its resemblance to substrate anteriority markers would be just accidental. This similarity would have played a crucial role in the emergence of the anterior category in Kriyol (Kihm 2011:97).

Another point, where the substrate seems to have had just a marginal importance, is the fact that Kriyol uses the verb *bin* ‘to come’ in order to express the specific future; *bin* is used in conjunction with the continuous aspect marker *na*, as we can see in the following example from Kriyol:

- 4) Kin kungsi si nomi, el ku na bin kasal.  
 who know her name 3sg.TOP that spec.fut. marry-3sg  
 ‘Whoever guesses her name will marry her.’ (Kihm 1994:108)

The use of *bin* as an auxiliary for the future probably comes from the Portuguese construction *vir a V*, lit. ‘go to V’. This Kriyol future form seems thus to be a semantic modification from the lexifier language, but we have to consider two more facts: first, such a future construction seems to be shared by a great number of creoles. Second, several Atlantic languages like Mandinka have the option of expressing posteriority by a periphrastic construction involving the verb ‘to come’. Since there seems to be no doubt that Kriyol *bin* derives from the Portuguese verb form *vem* (“s/he/it comes” < inf. *vir*), we have to conclude that substrate here has played just a “reinforcing” role and that “the ultimate force that drove the evolution was indeed the spontaneous reorganization of the TMA system of emerging Kriyol following a default blueprint” (Kihm 2011:99).

---

<sup>40</sup> See Lefebvre 1998.

A last example of the limited influence of the substrate in Kriyol is the negation *ka*: it comes from Portuguese *nunca* ‘never’ via lexical conflation. It is here that the substrate enters the stage: it is indeed unusual that the unstressed vowel is taken and not the stressed one; furthermore, Kriyol negation clitic *ka* keeps its meaning distinct from that of ‘no’, expressed by *nau*, whereas in Portuguese both meanings are conveyed by *não*. Evidence of the substrate influence is the fact that nearly all local Atlantic languages have a morpheme *ka* in their expression of negation.

Kihm’s conclusion is that the limited influence of the substrate depends on two important factors: first, Kriyol has emerged in a particularly rich linguistic environment, where no language was more important or diffused than others; that is why no local language has contributed consistently to the development of Kriyol, together with the fact that the initiators of Kriyol were perfectly bilingual.<sup>41</sup> Second, the limited influence of the substrate would depend on the status of the substrate itself: in Kihm’s words, its importance varies “according to the sociolinguistic setting of each creole emergence”. The scenario of Kriyol birth consisted of a group of individuals that were perfectly bilingual in Kriyol and another local language: their function was that of intermediaries between the local populations and the Portuguese. Their perfect bilingualism and the fact that they remained “a stable, endogenously reproducing group for a long time” and that their language was their only identity mark, made the substrate influence so limited, as we have seen above (Kihm 2011:101).

## 2.4 Kriyol grammar

So far, we have given an overview of the principal historical facts about the creole of Guinea-Bissau; furthermore, we have looked at a number of Kriyol substrate languages to see where this creole diverges and where it stays close to its substrate. Before we introduce the topic of the present study, namely the bare noun in Kriyol, we will look at its grammar and draw a sketch of its principal grammatical features.

Prior to starting our brief overview of Kriyol grammar, we should notice that there is no official orthography: instead of it, there is a proposal of official spelling in the *Projecto de*

---

<sup>41</sup> *Grumetes* and *filhos da terra*, the initiators of Kriyol, were perfectly bi- or multilingual. This means that they were able to separate the varieties they spoke and to limit the interferences between them (Kihm 2011).

*Ortografia e Separação das Palavras em Crioulo*<sup>42</sup>. However, it does not seem to have had any relevant effect on the orthographic behavior of comic books authors.<sup>43</sup> As a piece of evidence, during one of the sessions with my informants (Bologna, Winter 2009), all Kriyol native speakers, one of them wanted to show me how Kriyol is written: in order to do that, she took the Kriyol translation of the Bible as an example, but she was not sure and it seemed to me that she did not agree with the orthography used in that Bible.

#### 2.4.1 The morphosyntax of Kriyol

The creole of Guinea-Bissau is a language, which obligatory expresses its sentential subject.<sup>44</sup> The subject normally occupies the first position of the sentence, at the extreme left periphery. Sometimes this position can be occupied by an adverb, temporal or locative: in this case the subject follows the adverb.<sup>45</sup>

- 5) Li i ka ta fumadu.  
 there 3sg Neg HAB smoken  
 ‘You can’t smoke there.’
- 6) Amanha e na inaugura kil lossa.  
 tomorrow 3pl CONT inaugurate DEM shop  
 ‘That shop will be inaugurated tomorrow.’

Kriyol word order is quite rigid: it seems to be a characteristic common to creoles in general. The typical word order of a Kriyol simple sentence which contains a transitive verb is schematized in (7):

- 7) (Adverb) + Subject + (Negation) + (Aspect) + Verb + Object + (Tense)  
 +(Adverb)

<sup>42</sup> Portuguese for “Project of Ortography and Word-separation in Kriyol”.

<sup>43</sup> Comic books are the most productive (and almost unique) form of literature written in Kriyol in Guinea-Bissau.

<sup>44</sup> There seems to be (at least) one context, where the subject has not to be obligatorily expressed, i.e. with the verb *parsi*, “to seem”: *Parsi n kuma i para tchubi*, ‘It seems to me that it has stopped raining’.

<sup>45</sup> Sometimes an adverb can occupy the intermediary position between the subject and the verb:

*I presis tchiga sedu pa bu odja lugares.*  
 3sg be.necessary arrive early to 2sg find places  
 ‘You need to arrive early in order to get a seat.’

A crucial difference from other creoles is that GBC does not have the order Tense-Mood-Aspect: as we can see in (7), the aspect markers precede the tense markers. These latter ones, i.e. the tense markers, can be preverbal (e.g. *na+bin* for specific future) or postverbal (*ba*). Furthermore, in its usual postverbal position, *ba* also expresses the irrealis mood.<sup>46</sup> We will see these markers in more detail below.

The subject can be pronominal or lexical. In case of pronominal subjects, we may find both weak and strong pronouns or the weak pronoun only, i.e. the clitic.<sup>47</sup> Kriyol pronominal (clitic) system is schematized below in Table 1.

Table 1. *Kriyol pronominal system*

<b>Person</b>	<b>Full pronoun (optional)</b>	<b>Subject clitic (mandatory)</b>	<b>Direct/indirect object clitic</b>
<b>1sg</b>	<b>ami</b>	<b>n</b>	<b>n</b>
<b>2sg</b>	<b>abo</b>	<b>bu</b>	<b>u/bu</b>
<b>3sg masc./fem.</b>	<b>el</b>	<b>i</b>	<b>l/el</b>
<b>1pl</b>	<b>anos</b>	<b>no</b>	<b>nu</b>
<b>2pl</b>	<b>abos</b>	<b>bo</b>	<b>bos</b>
<b>3pl</b>	<b>elis</b>	<b>e</b>	<b>elis</b>

Although the clitic system of Kriyol is not as rich as the Romance ones,<sup>48</sup> it is worth noting that it has two specialized forms respectively for the subject clitic and object one, both direct and indirect. As we can see in Table 1, Kriyol pronominal subject consists of two pronouns: the so-called full (strong) pronoun and the clitic (or weak) one. The obligatory pronoun is the

<sup>46</sup> See Kihm (1994:83-124) for a more in-depth discussion of tense and aspect in GBC.

<sup>47</sup> The term clitic, from the Greek word for “leaning”, refers to any grammatical or lexical unit which forms a single accentual unit with the preceding or following word, depending on whether it is “proclitic” or “enclitic”. In the first case, the clitic precedes the word, whereas if it is enclitic, it will follow the word. See Matthews 2007:59.

<sup>48</sup> Romance languages have generally very rich and varied clitic systems, where there is a specialized form for almost every combination of gender and number. There are also more inflectional categories, say “cases”: subject (nominative), dative (recipient), accusative (patient), partitive and locative. The latter two have syncretic forms in languages such as Italian.

clitic one; the strong is optional and has topicalizing effects. This can be shown in the following examples:

- 8) (El) i findji ba durmi pa i ka perturbedu.  
 3sg.TOP 3sg pretend ANT sleep to 3sg NEG disturbed  
 ‘He was pretending to sleep in order not to be disturbed.’
- 9) (Abos) bo falal ba kuma i pudi ba bai ora si misti.  
 2pl.TOP 2pl say-3sg ANT that 3sg can ANT go now if want  
 ‘You told her that she was free to go, if she wanted to.’

Comparing the above sentences, we may note that the clitic subject is always used, whereas the full pronoun can be omitted; it is an individual choice of the speaker. In other words, the subject clitic can be “doubled” by the subject pronoun, which, on the contrary, is optional. As we already said, the realization of the subject in clitic form is one of the possibilities Kriyol disposes of in order to satisfy the EPP requirement.<sup>49</sup> The second possibility is to use a lexical subject, as we can see in the examples below:

- 10) Paulo tira si kazaku.  
 Paul take.off his/her jacket  
 ‘Paul took off his jacket.’
- 11) Paulo skirbi un karta pa João.  
 Paul write INDEF letter to John  
 ‘Paul wrote John a letter.’

The sentences in (10) and (11) have a lexical subject in the first position. Such a lexical subject could be substituted with a clitic subject and the resulting sentence would be grammatical as well.

Since the subject in Kriyol must be obligatorily always realized, the situation will not change in case of zerovalent verbs, like the following Kriyol example will show:

- 12) I na tchubi vs. \*na tchubi.  
 3sg CONT rain CONT rain  
 ‘It’s raining.’

---

<sup>49</sup> By EPP, or Extended Projection Principle, is meant the fact that the subject, even when expletive, must always be assigned to the verb.

This is the same situation we find in languages such as English, German or French: in the presence of a zerovalent verb like a meteorological one, they use an expletive subject,<sup>50</sup> which normally corresponds to the third singular clitic. Similarly to Kriyol, which uses third singular clitic *i* as the expletive subject, English adopts *it*, in German we will find *es* and in French *il*:

- 13) a. It rains vs. \*Rains. (Eng.)  
 b. Es regnet vs. \*Regnet. (Ger.)  
 c. Il pleut vs. \*Pleut. (Fr.)

On the contrary, in Romance languages such as Italian and European Portuguese, one does not find any expletive subject. These languages are pro-drop languages<sup>51</sup> and do not need any expletive subject, since they are said to have a phonologically empty subject position:

- 14) a. Piove vs. \*Esso piove. (It.)  
 b. Chove vs. \*Ele chove. (EP)

Worth noting is the split between first and second persons, on the one hand, and third person on the other. We may, indeed, easily note a different behavior of the clitics that lexicalize the P-property of Person from the clitics that, on the contrary, lexicalize the N-property of Noun. First and second persons are directly linked to the discourse event in the sense that their reference is immediately linked to the discourse world, as they respectively represent speaker and listener. First and second persons, both singular and plural, are always expressed by pronouns. The split mentioned above lies exactly in this fact, as the third person, both singular and plural, contrarily to the first and second ones, has a double choice: it can be expressed either by a clitic or lexical subject. In other words, the third person clitic subject is in complementary distribution with the lexical subject: it means that if there is a third person lexical subject in the sentence, the clitic cannot co-occur, and vice versa.

Kriyol clitic system has a specialized form for each combination of person and number, except for the third person, both singular and plural: here Kriyol shows syncretic forms for subject and object clitics. As represented in Table 1, the third person subject forms *el* and *elis*, respectively singular and plural, are syncretic with the (in)direct object clitic forms. As usual in the Kriyol clitic system, there is a contracted form *l* for the direct object, in case both the direct and indirect objects are in enclisis on the same verb.

<sup>50</sup> A subject is said to be “expletive” when it has no thematic role (Donati 2002:101).

<sup>51</sup> “Pro-drop” stays for “pronominal dropping”. The empty category where the phonologically null subject inserts is said “pro”.



As far as the clitic object, it always follows the verb; it means that the object is always in enclisis on the verb. Table (1) above shows that Kriyol has syncretic forms for indirect and direct objects, if only one of them is present in the sentence. On the contrary, if the sentence contains a trivalent verb<sup>52</sup> and both indirect and direct objects are expressed in the form of clitics, the indirect one will insert at the immediate right of the verb, between the verb and the accusative and will appear in its contracted form; on the other side, the direct object will appear in its non-contracted usual form. This is shown in (15):

- 15) I na dal el.  
 3sg CONT give-3sg 3sg  
 ‘S/he gives it to him/her.’

If the indirect object is clitic, it will insert at the immediate right of the verb, as we can see in the following sentence:

- 16) I dan libru.  
 3sg give-1sg book  
 ‘He gave me the book.’

If both objects are lexical, the accusative will immediately follow the verb and will thus precede the dative that on its turn will be introduced by the preposition *pa* ‘to’. We will have the prepositional dative also when the accusative is a clitic, which will be enclitic on the verb as usual.

- 17) I skirbi un karta pa Maria.  
 3sg write INDEF letter to Mary  
 ‘He wrote Mary a letter.’

#### 2.4.2 The nominal domain

One of the typical features of creoles is a reduced morphology, both inflectional and derivational. A direct consequence of this fact is the higher degree of morphosyntactic “analyticity” that creoles show if compared to their lexifiers. It implies that creoles with a reduced morphology normally have no (inflectional) distinction for gender, person and

---

<sup>52</sup> A verb is said “trivalent” when it needs three arguments in order to satisfy its valency: one external (the subject) and two internal (direct and indirect object) arguments.

number, as it happens in Kriyol. Typically, in Kriyol, and generally in creoles, nouns are bare: the noun phrase consists of the noun only. Verbs may be bare as well, whenever they are not accompanied by any tense or aspect markers. As for the Kriyol nominal system, it does not have a proper form of definite article. In contexts of determinate reference we may find the demonstrative adjective: *libru* ‘book’ vs. *kil libru* ‘that book’. On the other hand, Kriyol has a proper form for the indefinite article *un* ‘a, an’ and its plural counterpart *uns* ‘some’: these forms do not show any morphological change depending on gender conditions, since Kriyol has no morphological gender distinction. As for number, the quantification depends normally upon the context, but a morpheme *-s* (or *-is/-es*) sometimes adds to the noun, lexicalizing plurality: *karta/kartas* ‘letter/letters’.

### 2.4.3 The verbal domain

As we mentioned in section 2.5.2, a Kriyol verb may be bare just as the noun: the verb form does not show any agreement morphology with respect to person and number, nor does it have inflectional forms for the expression of tense and aspect. In other words, Kriyol does not dispose of morphological verbal inflection. The person and number distinction depends on the external argument of the verb, i.e. the subject: this could well explain the fact that the subject in Kriyol must always be expressed.

In Kriyol, most verbs that end in a vowel seem to belong to the same verbal classes as in Portuguese and the other Romance languages: *-a* (e.g. *fala* ‘speak, talk’ > EP *falar*; *tarbadja* ‘work’ > EP *trabalhar*), *-e* (e.g. *kume* ‘eat’ > EP *comer*) or *-i* (e.g. *obi* ‘hear’ > EP *ouvir*). There are in Kriyol also verbs that end in consonant like *ten* ‘have’ or in a vowel different from those of the Romance verbal classes: *sedu* ‘be’.

A direct consequence of the fact that the verb is bare is that it is often accompanied by aspectual or temporal markers. When it occurs bare, two different interpretations are possible depending on whether the verb is dynamic or stative: in the former case, we will have a simple perfective interpretation of the unmarked form of the verb, whereas if the verb is stative, the reading will denote a presently obtaining state of affairs (Kihm 2011:92). This translates into a past perfective reading and a simple present reading, respectively.

Nonperfective aspect and (nonperfective) past tense in GBC are overtly expressed by means of aspect and tense markers: *na* and *ta* are imperfective aspect markers, which occur at

the left of the verb verb. On the other hand, the past marker *ba* occurs at the right of the verb and, as remarked in Kihm (1994:104), it cannot insert between the verb and its enclitic object pronoun, if present. A further element indicating tense is the verb *bin*: whenever used in conjunction with *na* (*na + bin*), it acts as a preverbal tense marker for the expression of specific future (1994:108-113).<sup>53</sup>

Kriyol disposes, thus, of two imperfective aspectual markers, i.e. *na* und *ta*. The crucial difference between them lies in that *na* is used with eventive verbs and makes the interpretation continuous (either progressive or not). The event which the verb refers to is simultaneous to the reference and/or speech time. On the other hand, *ta* expresses a habitual or indefinitely reiterated process.

- 18) I na odjan.  
3sg CONT look-1sg  
'S/He is looking at me.'
- 19) I na dal el.  
3sg CONT give-3sg 3sg  
'S/He will give it to him/her.'
- 20) Li i ka ta fumadu.  
Loc 3sg Neg HAB smoke-ed  
'You can't smoke here.'

In the sentence in (20) we may also observe the negation clitic element *ka*, which always occurs at the left of the verb or, whenever an aspect marker precedes the verb, the negation will insert at the left of the aspect marker. The negation cannot, in fact, separate the verb from its aspect marker.

On the other hand, the imperfective anteriority is expressed via the tense marker *ba*, which is postponed to the verb:

- 21) Kasa ku kai bedju ba di mas.  
House that fall old ANT of very  
'The house that fell down was very old.'

---

<sup>53</sup> *Bin* can also act as a temporal adverb indicating posteriority. See Kihm (1994:108-113) for a more in-depth discussion.

- 22) El i yera ba un mediku ku djintis gosta ba del  
 3sg.TOP 3sg be.PAST ANT INDEF doctor that people-PL like ANT of-3sg  
 tchiu pabia i kura ba manga di djintis.  
 much because 3sg heal ANT lot of people-PL  
 ‘He was a doctor very appreciated by the people because he had healed a lot of  
 people.’

Another tense marker is lexicalized by the string *na + bin* and expresses the specific future. *Bin* has a double function in Kriyol grammar: it is both part of a posteriority tense marker and a verb with the meaning of ‘come’. For the interpretation of non-specific future Kriyol grammar supplies the construction *na + ba + ta*: it expresses events or states of things that are continuous or repeated in an extended future which starts at the speech time or when the circumstances are met.

- 23) Kin kungsi si nomi el ku na bin kasal.  
 Who know her name 3sg.TOP who spec.fut. marry-3sg  
 ‘Whoever knows her name will marry her.’ (Kihm 1994)

Another interesting feature of Kriyol grammar is the alternation between null and overt copulas. In adjectival predicates, the copula is normally not lexicalized:

- 24) Ami n kontenti.  
 1sg.TOP 1sg happy  
 ‘I’m happy.’

This depends on the fact that there seems to be in Guinea-Bissau Creole a class of verbal adjectives, which share some of the properties of Kriyol verbs: in fact, they can be accompanied by aspect and tense markers (Kihm 1994, 2007b). Worth noting is that Kriyol grammar is also endowed with overt copulas such as *sedu* ‘be’, *(y)era* ‘be.PAST’ and the third singular pronoun *i* which seems to act as a kind of pronominal copula. The overt copulas of Kriyol and the alternation overt/null copula will be dealt with in more detail in Chapter 4.

As far as it regards direct interrogative sentences, the word order is the same as in declaratives: there is no inversion between subject and verb; this resembles EP (and in general Romance) interrogatives and is, thus, different from languages such as English and German.

- 25) Bu nega si proposta?  
 2sg negate his/her propose  
 ‘Did you negate his/her propose?’

Direct interrogatives can be introduced by clusters of two interrogative elements and not just by a single element as in Germanic or Romance languages. In case of interrogative sentences such as *Kin ku/ki bin?* ‘Who came?’, *ku/ki* is a relative pronoun, the subject of *bin*, whereas *kin* is in focus. This would resemble a sentence like EP *Quem é que veio?* ‘Who (is it that) came?’ (Alain Kihm p.c.). In (27), we have a phonological contraction between *ku* in its role of complementizer and the subject clitic *e* ‘they’. (Kihm p.c.).

- 26) Kin ku/ki= who  
 Ke ku = what  
 Kuma ku = how  
 Kal...ku = which  
 Nunde...ku = where  
 Kantu...ku = how much  
 Kal ora...ku = when
- 27) Kal ora ke na djubi na spidju?  
 When 3pl CONT watch on mirror  
 ‘When do they look at themselves in the mirror?’

As far as it concerns another mood of the verb, the imperative, in its 2 singular person, it is lexicalized by the bare verb, whereas 1 and 2 plural persons are preceded by the corresponding subject clitics. In case of negative imperative, the negative clitic *ka* will insert before the subject clitic which, in its turn, will occupy the immediate left of the verb. First and second plural person of the imperative, both positive and negative, will be always preceded by the subject clitic:

- 28) Kuri! / Bo kuri!  
 run 2pl run  
 ‘Run! (You) run!’
- 29) Ka bu bai! / Ka bo bai!  
 Neg 2sg go / Neg 2pl go  
 ‘Don’t go!’

- 30) No bai!  
 1pl go  
 ‘Let’s go!’

As far as embedded sentences, there are in Kriyol two different complementizers of the kind ‘that’ which may introduce an embedded clause: *kuma* and *pa*. The former one, *kuma*, is used with epistemic and declarative verbs, whereas *pa* introduces embedded sentences containing factive verbs. The examples below will clarify the differences in the use of the two complementizers:

- 31) E pensa ba kuma e pudi ba diskansa mbokadu.  
 3pl think ANT COMP 3pl can ANT relax a.bit  
 ‘They thought they could chill out a bit.’
- 32) Ami n disidi pa fasi un bias lundju pa manhera ke  
 1sg.TOP 1sg decide to do INDEF travel long for manner that  
 ami n na bin sta manga de mis fora.  
 1sg.TOP 1sg spec.fut. stay lot of month outside  
 ‘I have decided to do a long travel, so that I will stay away for a lot.’

In the sentence in (31) the complementizer *kuma* introduces an embedded sentence, where the predicate is lexicalized by the epistemic verb *pensa*; on the contrary, the complementizer *pa* in (32) introduces the factive verb *fasi*.

Before concluding this brief review of the principal grammatical facts of Kriyol grammar, it is useful to describe some other important facts of the verbal domain. So far, we have focused on the active form of the verb; the formation of the reflexive should also be taken into account. Unlike in Romance or Germanic languages, Kriyol does not have any reflexive clitic pronoun, like for example Italian *si* or Portuguese *se*. Kriyol happens to be innovative if compared to its lexifier language: it uses for the reflexive a nominal phrase of the kind (*si*) *kabeza* ‘(3sg.POSS) head’, (*si*) *kurpu* ‘(3sg.POSS) body’; such expressions follow the verb phrase and occur, thus, in the typical position of the internal argument of the verb:

- 33) N ta laba kurpu.  
 1sg HAB wash body  
 ‘I wash myself.’

- 34) N ka na odia nha kabeza.  
 1sg NEG CONT hate my head  
 ‘I don’t hate myself.’

*Kabeza* and *kurpu* derive from the corresponding Portuguese nouns for ‘body’ and ‘head’, respectively *cabeça* and *corpo*.

On the other hand, the passive is built on a form derived from the Portuguese past participle with suffix *-du*: *pirdidu* ‘lost’, *pistadu* ‘lent’, etc. Passive construction assumes thus the form illustrated in (35):

- 35) subject + *i* (3sg) + past participle  
 36) Kil poesia i skribidu pa Djon.  
 that poetry 3sg written by John  
 ‘That poetry was written by John.’  
 37) El i pistadu dinieru.  
 3sg.TOP 3sg lent money  
 ‘Money was lent to him.’  
 38) Kil loja na bin inauguradu amanha.  
 DEM shop spec.fut. inaugurate-ed tomorrow  
 ‘That shop will be inaugurated tomorrow.’

The examples in (36) and (37) show the construction of the past participle above described. As for (37), we have to note that the verb *pista* ‘lend’ is trivalent, i.e. it needs three arguments: two internal arguments (the two objects of the verb, i.e. direct and indirect, or patient and recipient) and an external one (the subject of the predicate), in order to satisfy the EPP requirement. Here *pista* satisfies its valency by lexicalizing its internal objects in the same way the corresponding English verb ‘lend’ does: with a double accusative construction. This allows its passivization by the way described in the example in (37). As regards the sentence in (38), it contains a future passive, where the cluster *na + bin*, responsible for the specific future interpretation, is easy to recognize.

Another interesting verbal construction is the causative. In Kriyol there are two types of causative formation: the first one is similar to the English ‘to make + V’ (e.g. ‘to make cry’) and involves two verbs, of which the first one is the one which causes what is described by the second one. The verbs used for causativization are *manda* ‘send’ and *fasi* ‘do’. The following sentences will show these facts:

- 39) Mame manda mininus sai di kasa.  
 mother send child-pl go.out of house  
 ‘The mother makes the children go out of the house.’
- 40) N fasiu pirdi comboio.  
 1sg make-2sg miss train  
 ‘I have made you miss the train.’

The second possible causative form in Kriyol consists in adding the following suffixes to the verb: *-nta*, *-nte*, *-nti*; they are selected depending on the thematic vowel of the verb. Crucially, only a few verbs add such causative suffixes: e.g. *badja* ‘dance’ > *badjanta* ‘make dance’, *karga*, ‘load’ > *karganta*, ‘make load’, *kume*, ‘eat’ > *kumente*, ‘make eat’, *bibi* ‘drink’ > *bibinti* ‘make drink’, etc.

So far we have given an overview of Kriyol grammar. This will be useful for inserting in their natural ‘environment’ the facts which we will treat in Chapter 4, namely the distribution and interpretation of bare noun phrases in Guinea-Bissau Creole.



## Chapter 3

### Theoretical background

#### 3.1 Definiteness

In her analysis of the determiner system of Mauritian Creole, Guillemin (2011) takes definiteness, specificity and deixis as universal categories of meaning. Each of these categories has to be somehow expressed in every language. We can assume here a parametric perspective of the facts and claim that every language “chooses” how to express them.

This will be exactly the topic of the present chapter, which aims to provide an overview of the most important aspects that pertain to the level of the meaning, namely the categories of meaning introduced above: semantic categories such as definiteness, genericity and specificity will be taken into account. The conditions under which they are met will turn out to be crucial with respect to our analysis of Kriyol BNs in Chapter 4. Moreover, we are going to review in the present chapter some of the major approaches with respect to bare nominals. This will introduce us to the study of such noun phrases. Both the semantic and the syntactic perspectives will be taken into account.

To begin, we can firstly take into account the category of definiteness, which is realized via the definite determiner in a great number of languages. For example, Germanic and Romance languages have opted for the parametric variable described above, i.e. it is the definite article that expresses the category of definiteness in these languages. Furthermore, depending on the language, we can find the features ‘number’ and ‘gender’ marked on the determiner. More specifically, English uses the definite determiner *the*, which is a syncretic form both for singular and plural, masculine, feminine and neuter, whereas German has a specialized form for both gender and number. Furthermore, German expresses the case as well via morphology, thus deriving a specialized form for each combination of gender, number and case. As a result, we have some syncretic forms, e.g. *die* realizes nominative and accusative cases for the singular feminine as well as the plural forms of these cases for all three genders. The situation in Romance languages is very different from the German case: this is a direct consequence of the loss of the morphological case in Romance languages. At the same time, Romance languages present a more complex determiner system with respect to

English: for example, Italian has a specialized form for every combination of number + gender, deriving the forms *il*, *lo*, *la*, *i gli*, and *le*. *Il* and *lo* (for this latter form, in most cases, one finds in Italian the elided form *l'*) are different forms for the singular masculine definite determiner. The speaker cannot freely decide which one to use; the choice happens, in fact, on the basis of phonologically-driven factors such as whether a word begins with a vowel or with a consonant. In the first case, if the word is masculine, the speaker will use the elided form *l'*. *La* is the unique form for the feminine singular, whose plural counterpart is *le*, whereas *gli* is the plural counterpart of *il* and *lo*. Something similar happens in Portuguese, where the situation is however lexically simpler: *o*, *a*, *os* and *as* are the forms for the masculine and feminine singular and the masculine and feminine plural, respectively.

As we said at the beginning of this section, crosslinguistic variation shows that there are also languages that do not use determiners in order to express the (in)definiteness category: following Guillemín (2011:48), Mauritian Creole “does not overtly mark this contrast” and Kriyol does not do so either. It is one of the factors which often lead us to think that creoles have a sort of special grammar; it is important here to note that there are other non-creole languages that lack such an overt realization of the definite determiner as well. An example is Russian: like Mauritian and Guinea Bissau creoles, it has no (overt) determiner in order to express definiteness. Russian also lacks indefinite determiners. On the other hand, as we will see below, both MC and GBC have an indefinite determiner to lexically realize indefiniteness. Guillemín’s assumption about this fact is that MC has a phonologically null definite determiner. We will discuss this assumption in the next section, where the syntax of bare nouns will be taken into account. For the time being, we will give a description of the semantic category of definiteness.

Before I draw a brief historical sketch of theories about definiteness, let me clarify that the option of non-overtly marking the definite character of a noun phrase is an option which is largely present in languages that overtly realize the definite determiner. This explains why we have bare nouns also in languages such as English or German, and not only in languages that lack an overt definite determiner. In her analysis of proper names, Matushansky claims that according to Stowell (1989), “verbs of nomination can appear with a bare nominal predicate” (Matushansky 2005:4) as the following example shows:

- 41) The queen appointed her lover treasurer of the realm.

The same is true for naming construction like (42):

42) Anne's death made George (the) king of England.

(Matushansky 2005)

Both constructions admit the omission of the definite determiner. As claimed by Matushansky the fundamental condition for the determiner's omission is that there is only one individual satisfying the predicate (at the relevant moment). In such constructions, the determiner is not present in languages that require an overt determiner in the argument position; some of the languages that show this phenomenon are the Uzo-Aztec language Pima, some Northern Italian dialects, European Portuguese and Northern Norwegian, to mention some of the languages taken into account by Matushansky. The following example is from Northern Norwegian, a language which requires a definite determiner introducing a noun phrase in the argument position:

43) Han heter (\*han) Øystein.  
he is-called he Øystein  
'He is called Øystein.'

The grammaticality of the sentence in (43) depends on the absence of the definite determiner *han*. On the contrary, if it is present, thus introducing the proper name Øystein, the sentence would result ungrammatical. According to Matushansky, the facts described so far would not depend on some special properties of the determiner, but on the predicate position these bare definites occupy in naming constructions and with nomination verbs.

### 3.1.1 Approaches to definiteness

In the course of the last century, several scholars have tried to give a definition of (in)definiteness. Most approaches are based on the concepts of *uniqueness*, *existence* and *familiarity*. Guillemin (2011) briefly sketches Russell's (1905) theory, which speaks of "definite descriptions", whose use is conveyed by the conditions of *existential commitment* and *uniqueness*. This type of analysis can be applied only to definite singulars, which really show the two above mentioned requirements. On the contrary, this kind of "quantificational analysis" cannot be applied to the analysis of definite plurals: it is in fact not possible to substitute the plural definite determiner with quantifiers like *all* or *every*, at least not in every context. As explained in Guillemin, it seems to depend on the fact that "like singular definite descriptions, plural definite descriptions are dependent on a particular context for

interpretation. The quantificational analysis of *the* thus predicts the ‘Familiarity’ effect in the sense that it expresses a proportion of a set that the hearer must be able to identify for clear interpretation” (2011:50; emphasis in original).

We can now introduce the traditional “Theory of Familiarity” by Christophersen (1939), who aims at completing the analysis of Russell in a different way. According to Christophersen, the definite article is used in order to speak of a referent which is already known in the discourse, whereas the indefinite determiner introduces new information. The speaker is always supposed to have in mind the referent, if one uses the definite determiner, i.e. the referent has to be “familiar”. The referent already known in the discourse world of the speaker and hearer can be either supplied in the form of information in the discourse or via shared knowledge. A consequence of that, as shown in Guillemin (2011), is that, whilst definites must have antecedents in the discourse, indefinites do not have any, exactly because of the fact that they introduce new information.

In more recent times, the terminology of Christophersen, i.e. the term *referent*, has been substituted for the sake of major clarity, by *discourse referent* (Karttunen 1971, Heim 1983, 1988). As explained in Guillemin (2011), a discourse referent is not always a unique referent, contrary to Russell (1905); this explains the fact that we sometimes meet definite determiner phrases that do not refer. Consider the following example from Christophersen (1939:140 in Guillemin 2011:49):

44) Towards evening we came to the bank of the river.

The discourse referent introduced by the definite determiner is *bank (of the river)*, and it fails to refer with respect to the fact that every river normally has two banks. On the basis of the work of Christophersen, Hawkins (1978) lists the uses of the definite article as follows:

1. direct anaphora
2. associative anaphora
3. visible situation use
4. immediate situation use
5. larger situation use (specific knowledge about the referent)
6. larger situation use (general knowledge about the referent)
7. unidentifiable uses

The first type of use of definite noun phrases, the direct anaphora, makes reference to the fact that the definite noun phrase has been licensed before in the discourse by the same noun

phrases introduced by an indefinite determiner. Languages like English use in these cases a definite determiner, whereas Mauritian Creole, as discussed in Guillemin (2011), and Kriyol do not.

Moreover, by associative anaphora it is meant that the definite noun phrase is not licensed by the very same noun phrase introduced by an indefinite determiner, but by another indefinite noun phrase, whose noun is semantically and contextually associated with the definite noun phrase in question.

In the third type of use, the visible situation use, the definite determiner could often be substituted by a demonstrative, but when the object is not visible, the substitution by demonstratives cannot take place. On the other hand, examples of larger situational uses are “The Queen of England” or unique nouns like “the sun”.

As for the last type of use of definite descriptions, the unfamiliarity uses, Guillemin assumes that they contain “a modifying noun phrase, a relative clause or a proportional phrase, which serves to establish a definite referent for the hearer without the need for previous mention” (2011:54) or when the hearer cannot identify the object, when there is no presumption of prior knowledge, and no associative link between the referent and the other previously mentioned individual.

Following Guillemin (2011), Russell (1905) also claims that singular definite descriptions contain an assertion of existence. According to Vendler (1971), Russell’s claim is not completely true, more precisely it does contain a true generalization, i.e. that singular definite noun phrases must be “licensed” in terms of existential assertion, but the “locus” for the occurrence of this sort of licensing is not the definite noun phrase, but another sentence, without which it would not be possible for the noun phrase to be definite. As exemplified in Vendler (1971, in Guillemin 2011:50): “The bear I shot yesterday” must have a presupposed shared knowledge that “I shot a bear yesterday”. Such sentences, which need to be “licensed” on the basis of a previous (often omitted) sentence, are defined by Strawson (1950) “incomplete definite descriptions”, whereas those sentences that do not depend on the context, are defined as “complete definite descriptions”. These latter ones refer to unique individuals taking no care of the context, e.g. “the author of Waverley”.

With respect to the above mentioned features of *uniqueness* and *familiarity*, we have to introduce another account on definiteness worked out in Farkas and de Swart (2007). They analyse an additional feature that helps determine the contrast between definiteness and

indefiniteness, namely *maximality*, which amounts to *uniqueness* if we are in the presence of a plural noun phrase.<sup>54</sup> Such features, *uniqueness/maximality* and *familiarity*, concern the discourse referent as defined above and represented by the noun phrase, and are considered as parameters with respect to which a noun phrase can be [ $\pm$  max], [ $\pm$  fam] depending on whether the noun contains these features. A noun phrase which encodes [+ max] has a unique referent if singular, whereas if it is marked as plural, it has maximal reference. Furthermore, if a noun phrase is familiar in the relevant situation, “or accomodatable therein” (Farkas and de Swart 2007:9), it is [+ fam]. In cases where both maximality and familiarity are encoded in the noun phrase in question, Farkas and de Swart (2007) propose the term *dynamic reference* from Farkas (2002). In their analysis within the framewok of Optimality Theory, they also claim that the unmarked option with respect to the contrast between definiteness and indefiniteness is the latter one: “definite forms impose a semantic requirement with respect to which indefinite forms are neutral” (Farkas and de Swart 2007:10).<sup>55</sup> This is a so-called “asymmetric” account of the contrast definiteness vs. indefiniteness, where indefinites are normally characterized by the absence of uniqueness and familiarity. On the contrary, a “symmetric” account of the facts would consider both ends of the contrast as characterized by the requirements of maximality and familiarity. Indefinite nouns do not require any restriction on familiarity or maximality conditions, as opposed to the situation of definite noun phrases, which must be familiar, as we already said above, and in the case of bare nouns, the preferred interpretation is the definite one if the noun is both maximal and familiar. If the noun phrase in question is inserted in an episodic sentence, we have a definite article in the case of uniqueness and familiarity, whereas an indefinite determiner is present if these properties do not belong to the context in question, i.e. if the discourse referent is [- fam] and [- max].

### 3.2 Genericity

When we speak of genericity, we have to consider that there are traditionally two kinds of genericity, as discussed in Krifka (1995). The first type regards nouns and confers them the skill of referring to a kind rather than to a particular object. On the other hand, the second type does not relate to a noun phrase only, but rather to an entire sentence, namely a

---

<sup>54</sup> The crucial role played by “maximality” with respect to the issue of definiteness has been recognized before Farkas & Swart (2007), e.g. by Link (1983). Manfred Krifka (p. c.).

<sup>55</sup>This has also been stated before Farkas & Swart (2007), e.g. by Heim (1987). Manfred Krifka (p.c.).

generic sentence. Krifka refers to the latter as *characterizing sentences* that describe regularities of behaviors in opposition to *particular sentences* “which express statements about particular events, properties of particular objects, and the like” (Krifka 1995:3).

Consider the following sentences from Krifka (1995):<sup>56</sup>

45) The potato was first cultivated in South America.

46) John smokes a cigar after dinner.

The sentences above are respectively examples of *kind-referring* or simply *generic* noun phrase and *characterizing* or *generic sentence*, as discussed above. In (45) the *potato* does not refer to a specific potato, but to the kind itself, and is thus a generic noun. On the other hand, in (46) the whole sentence is a generic, say characterizing one, insofar as it describes one of John’s regular behaviors in his smoking a cigar after dinner. These two kinds of genericity may intersect giving rise to sentences of the following type:

47) The potato is highly digestible.

Like the one in (45), the *potato* in (47) is a generic noun insofar as it does not refer to a specific potato, but to the kind “potato”. Furthermore, (47) is a generic sentence since it expresses a regularity of the behavior of the tuber under consideration in its being highly digestible. Following Krifka, characterizing sentences may allow for exceptions unless they lose their truth value: this is not true in case of universally quantified sentences<sup>57</sup> that “make a claim for every object of a certain sort” (1995:4). The sentence in (46), repeated in (48) below, can help us understanding that:

48) John smokes a cigar after dinner.

The sentence above is not a universally quantified sentence, but simply a characterizing sentence that does not represent a “timeless truth”, i.e. a statement that is always valid. In fact, “it is perfectly possible to claim that a characterizing property held in the past or will hold in the future, without any implication for the present” (Krifka 1995:6).

---

<sup>56</sup> If not otherwise specified, the example from English and Chinese that I will use in this section to explain the facts being described here are from Krifka (1995).

<sup>57</sup> In predicate logic, by universal quantification is meant that something is true for everything or every relevant thing. The symbol for universal quantification is a turned ‘a’,  $\forall$ .

### 3.2.1 Generic sentences

The eventuality of facing a characterizing sentence does not properly depend on what kind of noun we use, e.g. proper names, definite or indefinite singular NPs, quantified NPs, bare singular or plural NPs. Furthermore, these types of sentences are often ambiguous, thus not always easily recognizable as characterizing. In the individuation of generic sentences we may clearly note the interplay of factors like the presence of temporal adverbs like *usually*, auxiliary constructions like *used to* or agentive nouns, to mention some of them. On the other hand, progressive and perfect sentences give rise to episodic, particular sentences, whereas if the sentence contains a present, past or future tense without any other “specific element” that enforces the one reading or the other, it may vary between a characterizing and a particular reading.

Following Krifka, one of the main claims about generic sentences is that they may be habitual or lexical. On the one hand, habitual characterizing sentences present verbs which are normally used to form episodic sentences (e.g. *smoke*). On the other hand, lexical characterizing sentences employ verbs which cannot raise an episodic reading (e.g. *know*). In other words, lexical characterizing sentences generalize about the properties of individuals, whereas habitual sentences generalize about events (Krifka 1995:17).

As for particular sentences, we have to distinguish between stative and dynamic sentences: the former describes properties of the subject, whereas the latter describes events which involve it. Worth noting is that all characterizing sentences are stative, whereas only some particular sentences are stative (e.g. *Simba is in the cage*). Before we continue with the description of the different types of sentences, let me briefly explain the characteristics of a characterizing sentence.

### 3.2.2 What makes a sentence (non)generic

As we said above, Krifka (1995) lists several grammatical elements which can yield a characterizing reading of a sentence. A first option is that we find a temporal adverb in the sentence: the presence of adverbs such as *usually*, *typically*, *always*, *often*, *sometimes*, etc., yields a characterizing reading. This is true also in sentences containing a progressive:

49) a. John is smoking a pipe.



- b. John is often smoking a pipe.

What we claim is immediately recognizable by comparing the two sentences in (49) above: both contain a progressive, but their interpretations are quite different with respect to the (non)genericity factor. More precisely, (49a) receives an episodic, particular reading because what is being predicated refers to the moment of speech only. In contrast, (49b) is a characterizing sentence insofar as the temporal adverb *often* is present, whereby the predication results into a habit. However, this would be true also if one substitutes the adverb *often* with *usually*, *always*, etc.

A second factor enforcing the characterizing reading of a sentence is the use of auxiliary constructions of the type *used to*. English is not the only language that shows such a construction; the German counterpart of the English *used to* is *pflegt zu*, and Italian has a similar construction, i.e. *essere solito* ‘be used (to)’, to mention just a couple of examples of the presence of such constructions in other languages. Like in English, the expression *essere solito* is sensitive to the agreement rules of Italian. In more detail, the verb *essere* ‘to be’ has to agree with the subject as to the feature person. On the other hand, distinct from English, Italian requires the past participle *solito* to agree with the feature gender (m. *solito* vs. f. *solita*).<sup>58</sup>

- 50) Ero solito/a andare a scuola a piedi.  
‘I used to go to school by foot.’

We now resume our description of factors enforcing the generic reading. A third factor is the presence of verb phrases in the middle voice, which is not an exclusive phenomenon of English. In fact, one may find this in Italian as well:

- 51) a. This shirt washes easily.  
b. Questa maglia si lava facilmente.  
this shirt 3sg wash easily

In (51a) the verb *wash* is used in the middle voice; in order to do that, English does not need to add any syntactic or lexical element. Distinctly, the Italian counterexample in (51b) shows that the third singular person clitic *si* is necessary in order to construct the reflexive. In the absence of this clitic, in fact, the sentence would result in the active voice and would give rise to an odd reading, where *maglia* would result in the subject of an active transitive verb. It is

---

<sup>58</sup> Worth noting is that this construction is perceived as “high standard” from Italian native speakers. In the everyday use of Italian, speakers preferably use sentences with temporal adverbs like *di solito*, “usually”, *spesso*, “often”, etc.

worth noting that the Italian clitic *si* covers more than one function: it can be reflexive for the third person, both singular and plural, and subject in impersonal constructions.

Other factors which may yield the generic reading are the use of agentive nouns like *smoker*, derivation of deverbal adjectives (Eng. *-able*, It. *-abile/-ibile*, Ger. *-bar*), some special lexical items with the function of describing generalizations (e.g. *to have an inclination to/the habit of*). Furthermore there is in some languages a specialized verbal morphology in characterizing sentences like the verbal prefix *hu-* in Swahili. In contrast, the types of noun phrases which can be found in a characterizing sentence are not limited, as we will see below.

### 3.2.3 Generic nouns

We have spoken so far of genericity at the sentential level, and must now introduce the matter of generic noun phrases, which is of particular significance in cases of bare nouns. Krifka (1995) defines generic nouns as “kind-referring”. In fact, they refer to kinds, and not to individual objects (or sets of objects) or, to say it in Krifka’s words, they “actually do refer to kinds, which are modeled as special types of individuals” (1995:64).

A kind-referring noun may take different forms. For instance, in English it can be a count or a mass noun, a definite singular, a bare plural or a bare mass noun. It is worth noting that the indefinite singular noun phrase cannot be properly considered as a generic noun along with bare plurals. It yields a kind-referring reading when it occurs in a characterizing sentence, but the “locus of genericity” is the sentence itself and not the indefinite singular noun phrase. As for bare plurals, although they are often seen as plural counterparts of singular indefinites, they can also denote a kind, in contrast to singular indefinites. In other words, bare plural NPs are normally interpreted as indefinites, just as well as their singular counterparts, whereas their definite, kind-referring interpretation is more suitable in syntactic environments such as categorical sentences and with stative verbs. In these contexts, crucial for the definite interpretation is also the position where the bare plural shows up in the sentence. In more detail, the definite reading of a kind-referring noun phrase is preferably derived when the noun is the subject of a categorical sentence (in topic position and with a stronger stress) or the object of a stative verb: “these verbs favor the definite interpretation of bare plurals” (Krifka 1995:73). As Krifka explains, these two positions are not properly suitable for introducing new information, normally interpreted as indefinite.

As for the kinds of noun phrases we may interpret as kind-referring, one has to note that in Italian and in Romance languages more generally, the situation is a bit different from that described above, insofar as the determiner system of Italian differs in some important respects from the English one. First of all, Italian does not allow bare arguments (Longobardi 1994, Chierchia 1998). In more detail, a bare noun, either singular or plural, cannot be a kind-referring expression: in order to get a kind-referring reading of an Italian noun, the determiner has to introduce the noun. In other words, “a ‘nominal expression’ is an argument only if it is introduced by a lexically filled D position” (Longobardi 1994:613, emphasis in original). We will take into account the syntactic structure of bare nouns later on in the present chapter. For the time being, we will just mention some exceptions noted by Longobardi itself. Indeed, Italian may have bare nouns in the argument position, but under more restricted conditions than in English, namely as a singular mass noun, plural count noun (bare plural) or rarely as a singular count noun in the scope of a sentential negation (Longobardi 1994:613). The facts described here are shown in the following examples from Longobardi (1994), which contain a singular mass, a plural count noun and a singular count noun in a negative sentence respectively:

- 52) a. Bevo sempre vino.  
       ‘I always drink wine.’  
       b. Mangio patate.  
       ‘I eat/am eating potatoes.’  
       c. Non c’era studente in giro. (Benincá 1980)  
       ‘There wasn’t student around.’

As we said above, the bare singular count noun is allowed only if a negation is present. If the sentence had a positive polarity, one should use a plural-marked form of the noun phrase. In order to maintain the quantitatively unspecified nominal expression, the positive counterpart of (52c) would be then *C’erano studenti in giro*, where the noun phrase is pluralized.

Take now a look at the following examples from English and Italian:

- 53) a. The lion is a predatory cat.  
       b. Il leone é un felino predatore.  
       54) a. Lions are predatory cats.  
       b. \*Leoni sono (dei) felini predatori.  
       c. I leoni sono (dei) felini predatori.  
       55) a. Gold is a precious metal.

- b. \*Oro é un metallo prezioso.
- c. L'oro é un metallo prezioso.

In the examples above, both English and Italian sentences contain kind-referring nouns: in (53) we have definite singular nouns in the subject position and indefinite singular nouns in the predicative position. In (54) the English subject *lions* occurs without determiner, it is thus a bare plural. In Italian this is not possible for the subject position, where the plural subject must be introduced by a determiner just as well as the singular subject in (53). In other words, Italian always needs a determiner in order to give rise to an argument, and this is the situation of Romance languages in general, as described in Chierchia (1998). Furthermore, Italian can optionally add the partitive article; this would introduce the noun in the subject position since Italian does not allow bare subjects. As for the sentences in (55), it is true what we have said so far for (54): Italian does not allow a bare noun in the subject position, even if it is a mass noun.

### 3.2.4 Taxonomic kind-denoting noun phrases

As noted in Krifka (1995), there is a “secondary count noun reading” that needs to be taken into account, namely the taxonomic reading. Here count and mass nouns can denote subspecies in a taxonomic hierarchy (1995:5) and have to be analyzed like “kind-denoting”.

- 56) One metal, namely copper, went strongly up on the market yesterday.

As kind-denoting we may find the same nominal variety we found above in case of kind-referring noun phrases: definite or indefinite singular count nouns, bare plurals and bare mass nouns. The same is true for object-referring nouns. There are just a few special cases of nouns or idiosyncratic expressions which happen to be always interpreted as kind-referring: examples from English are *man* and nominal expressions of the sort *this kind of/each species of* + N. All other nouns seem to vary from one type of reading to the next. For all other cases, Krifka (1995) reports some strategies in order to disambiguate among kinds of nouns. From these tests, it immediately results that only kinds can be the respective subject argument of verbs like *die out* or *be extinct* and object of *invent* or *exterminate*. There are also verbal predicates such as *be a mammal* or *be domesticated*, which can choose between two different kinds of subjects: the kind-referring, which is preferred if the subject is a general term, and the proper name. Krifka defines as *kind predicates* the verbs which give preferential rise to a kind-referring reading. It is worth noting that the verb phrase does not need to be stative:

indeed, one may also say *The panda is dying out* (Krifka 1995:63). Another point to take into account is that “it is not possible to form kind-referring NPs with just any nominal constituent. [...] the noun or complex nominal constituent must be semantically connected with a ‘well-established kind’ to which the noun phrase then can refer” (1995:11, emphasis in original). If not, kind-referring noun phrases (bare plural and bare singular noun phrases) might get an object-referring reading also in characterizing sentences, thus they behave in these cases like singular indefinites. In order to better understand that, consider the following examples from Carlson (1977) as reported in Krifka (1995):

- 57) a. The Coke bottle has a narrow neck.
- b. ?? The green bottle has a narrow neck.

The sentence (57a) is about a bottle of Coke, which is a well-established kind of bottle; that is why *the Coke bottle* receives a kind-referring reading, whereas the green bottle in (57b) is quite odd in a kind-referring reading. Therefore, it should receive an object-referring interpretation. All the observations on kind-referring noun phrases above suggest that the genericity of these nouns is tied to the noun phrase itself and not to the sentence (1995:63). The fact noted above that verbal predicates like *be extinct* and *invent* require a kind-referring noun, does not contradict the conclusion that the genericity is tied to the noun. Following Krifka, these verbs exert some selectional restrictions. As every count noun, taxonomic kind-referring noun phrases can display the form of every type of noun phrases: indefinite or definite, singular, plural or mass, bare plurals, NPs with numerals and quantified NPs. This similarity of behavior to count nouns may be observed also in several contexts, like the following examples from Krifka can show:

- 58) Two red wines are produced in Württemberg: Lemberger and Trollinger.
- 59) Chlorine is a halogen.
- 60) ?A dodo is extinct.
- 61) This kind of whale lives mainly in Arctic and Antarctic waters.

The sentence in (58) shows that a mass noun in a count noun context gives rise to a taxonomic reading. The same happens in (59), where a count noun like *halogen* can only be interpreted as taxonomic. In (60) we may observe that the taxonomic reading yields quite odd results because *dodo* is an empty set now. Finally, in (61) we do not have a simple noun phrase, but a construction introduced by *kind* (or *type*) which denotes subkinds. The same happens in Italian with constructions of the sort *tipo di* + N, whereas German has specialized suffixes to derive taxonomic count nouns: N + *-art/-sorte*. A very different way to provide a noun with a

taxonomic reading may be found in the classifier language Chinese, which has specialized classifiers for the taxonomic case. Take a look at the following examples:

- 62) Yi zhī xióng.  
one CL bear  
'An individual bear.'
- 63) Yi zhōng xióng.  
one CL bear  
'A kind of bear, a bear species.'

The sentence in (62) contains a count noun classifier *zhī*, which is different from the taxonomic classifier *zhōng* in (63).

It is important in this regard to define the nature of the subkind relation: here we have to speak of taxonomic hierarchies that precede hierarchically from kinds to subkinds and so on until one reaches the basic-level terms or level of everyday words that are used to describe objects and no more in order to speak of kinds. Such hierarchies can be represented via T, the taxonomic subkind relation: with the formula  $T(x,y)$  we mean that x is a subkind of y. The taxonomic subkind relation shows similarities to the realization relation R. Both are related via the following formula:  $[T(x,y) \ \& \ R(z,y)] \Rightarrow R(z,y)$ . This means that they “actually do refer to kinds, which are modeled as special types of individuals” (1995:64). As such, “if x is a subspecies of y and z is an object of y, then z also belongs to y” (1995:77). But the two relations T and R must not be confused insofar as the first relates kinds to kinds, and the latter objects to kinds.

### 3.2.5 Nontaxonomic kind-referring noun phrases

We claimed above that kinds are individual entities, which kind-referring nouns refer to. Following Krifka, these nouns should be semantically analyzed as proper names on the basis of some striking similarities between the two: both kind-referring nouns and proper names are definite, referring to expressions and, to the former, there belong some proper names as well. The question here is why the kind-referring NPs, which are not proper names, take the exact form they have, namely definite singular NPs, bare plurals, bare mass or proper names. Put in other words, there is a need to define whether the argumental (kind-referring) use of the noun is the primitive and consequently the predicative use is derived from it or vice versa. Or are both primitive uses?

A first answer can be found in Krifka (1995), who claims that in many languages common nouns are primarily predicates, although he does not exclude that there certainly are counterexamples. A more rigid conclusion, which has generated several critiques and further studies, is arrived at by Chierchia (1998). He claims that there are three types of languages, classified on the basis of the nominal primitive they present: argumental (kind-referring), predicative or both. We will see that in more detail below. For the purpose of this present section, we will limit our discussion to the answer of Krifka; he claims that “a common noun like panda has two functions: first, it is related to a kind, in this case *Ailuropoda melanoleuca*, and second, it is related to a set (or property), in this case  $\lambda x[R(x, \text{Ailuropoda melanoleuca})]$  – or the set of all pandas” (1995:66). Here R is the realization relation that relates kinds to their specimens:  $R(x, k)$  means that the object x belongs to the kind k. Let us designate a common noun as  $\delta$ . As we said above,  $\delta$  has two functions, a kind-referring and a predicative one; we can label the former as  $\delta_k$  and the latter as  $\delta_p$ . As Krifka claims, according to its kind-referring use,  $\delta$  behaves like a proper name, namely NP; whereas according to its predicative use, it behaves like a nominal predicate, i.e. N.

Several scholars have tried to confront these facts in more or less different fashions. For example, Quine (1960) claims that mass nouns are singulars if they occur in the subject position, whereas their presence in the postcopular position would give rise to a general reading. As we said above, the question is which one is the primitive use between the argumental and the predicative function. We will discuss a possible solution to this puzzle in Chapter 6. For the time being, it is sufficient to mention that our analysis will follow Depréz’s (2007) model, assuming every bare noun to be of the type k(ind). As a consequence, both kind-referring and predicative uses of BNs are yielded through derivation rules.

### 3.3 Specificity

It is time here to introduce an important difference among noun phrases as to their reference, namely noun phrases, which refer to a particular entity and noun phrases which do not. They are called respectively *specific* and *nonspecific*. *Specific* are thus NPs which refer to a particular individual or set of individuals in contrast to *nonspecific* NPs which do not have such a reference. This distinction does not depend on the distinction *kind-referring* vs. *object-referring*. Indeed, when we talk of kinds, we can of course do that either specifically or

nonspecifically. If we do not refer to any particular individual, i.e. nonspecifically, we will use an indefinite NP, which designates “some element in this kind taxonomy”.

About the contrast specific vs. nonspecific, Haspelmath (1997:37) assumes that “in some languages, different indefinite series are used depending on whether the indefinite NP is specific or non-specific”. There is no “universal” agreement about specificity. An explanation could be based on the presupposedness of “a uniquely identified” entity of which we are speaking, i.e. specific and, on the other hand, for a nonspecific reading the identity of the individual in question has not been established. So, following Haspelmath, “an expression is specific if the speaker presupposes the existence and unique identifiability of its referent” (ibid.). Haspelmath gives the example below to show the two possible readings, namely the specific (64a) and nonspecific (64b):

- 64) Nobuko wants to marry a native speaker of Ainu.
- a. Specific: she fell in love with him during fieldwork sessions.
  - b. Nonspecific: because she is Ainu herself, and she wants her children to acquire her ancestors’ language.

With further regard to specificity, there are three crucial correlates to the distinction between specificity and nonspecificity, i.e. discourse referent, paraphrasability by an existential sentence, and disambiguation by specific determiners (Haspelmath 1997:38). More specifically, “only a specific NP can have a ‘discourse referent’, i.e. can be referred to by an anaphoric pronoun in a present indicative clause” (Karttunen 1976:366).

- 65) Cheobai bought a bicycle (specific). It is black.

Similarly, a paraphrase with an existential sentence is only possible with a specific NP (Heringer 1969:90). The paraphrase for the specific reading in (64) above would be as follows (66):

- 66) There is a native speaker of Ainu who Nobuko wants to marry....

Last, the specific reading seems to be necessarily yielded by certain determiner-like expressions in some languages such as English ‘a certain’ or German *ein bestimmter* ‘a certain’:

- 67) Nobuko wants to marry a certain native speaker of Ainu.



A similar treatment of specificity is found in Guillemain (2001) and Farkas (1994). More specifically, as claimed in Guillemain (2011), specificity is one of the semantic universals which must find expression in the grammar of a language. Specificity intersects both genericity, (in)definiteness and number: noun phrases defined as [+ spec] can be kind-referring or object-referring, definite or indefinite, singular or plural. *Specific* are said to be noun phrases that refer to a particular entity, whereas *nonspecific* nouns do not. *Specific* nouns are thus NPs which refer to a particular individual or set of individuals in contrast to *nonspecific* NPs which do not have such a reference. These terms correspond respectively to *transparent* and *opaque* in the terminology developed in Quine (1960). As Krifka (1995) explains, this distinction does not depend upon the distinction kind-referring vs. object-referring. When we talk of kinds, in fact, we can of course do that either specifically or nonspecifically. If we do not refer to any particular individual, i.e. nonspecifically, we will use an indefinite NP, which designates “some element in this kind taxonomy”.

Just as well as the semantic universal of definiteness, specificity can be overtly marked on nouns or not. Following Guillemain (2011), specificity is not overtly marked on definite noun phrases in Germanic and Romance languages, whereas they can express it on indefinite nouns via the indefinite article; however, this could trigger a certain degree of ambiguity between a specific and a nonspecific reading.<sup>59</sup> The ambiguity between specific or nonspecific indefinite interpretations can be better understood if we consider the two interpretations of the following sentence:

68) Paul wants to buy a monkey.

The internal object of the verb phrase, *a monkey*, could be specific or nonspecific. In the first case, if it yields a specific reading, then the monkey has a referent in the discourse. On the other hand, if the monkey is a nonspecific noun phrase, it has no particular referent, and the noun phrase represents something new in the discourse. Guillemain gives to these two readings the following logical representation, which helps disambiguating:

---

<sup>59</sup> Guillemain (2011) claims that this ambiguity between a specific and nonspecific reading arises when the noun phrase in question is the complement of opacity inducing predicates such as intensional transitive verbs like ‘desire’, ‘need’ or ‘want’. She further defines intensional verbs as predicates, which are anomalous in at least one of the following three ways: “(i) interchanging expressions in the complement referring to the same entity can change the truth-value of the sentence embedding the VP; (ii) the VP admits of a special ‘unspecific’ reading if it contains a quantifier, or a certain type of quantifier; and (iii) the normal existential commitments of names and existential quantifiers in the complement are suspended even when the embedding sentence is negation-free” (Guillemain 2011:56).

69) Specific

- a. Paul wants to buy [*a monkey*]<sub>i</sub>. He saw it<sub>i</sub> at the market yesterday.  
 b.  $\exists x[M(x) \wedge \text{WANT}(p, \text{BUY}(p, x))]$

The formula in (69b) is adapted from Guillemin (2011); she explains that in this case MONKEY has scope over WANT and the specific indefinite *a monkey* can be substituted by a pronoun. In contrast, it is not possible in the nonspecific reading of *a monkey* as the representation in (70) can show, where WANT has scope over MONKEY:

70) Nonspecific

- Paul wants to buy *a monkey*. He will look for one at the market.  
 WANT (p, [Ex (MONKEY (x) & BUY (p, x))])

With respect to the terminology developed in Farkas (1994), the sentences above are examples of *scopal specificity*: the specific or nonspecific reading of the sentence in (68) depends on whether the indefinite is in the scope of some quantifier or intensional predicate, or not. In more detail, if a NP takes wide scope, it will be specific. This is exemplified in (69) above. On the other hand, a NP which takes narrow scope like in (70) will be nonspecific.

Under certain conditions, specific indefinites can be introduced by the demonstrative. In fact, “when the individual in question (i) has some ‘notable’ characteristic, or (ii) will be the focus of the forthcoming discourse” (Guillemin 2011:56-57), the demonstrative introduces the noun phrase. Coming back to definite noun phrases, there are also languages that mark it morphologically. For instance, Mauritian Creole uses the post nominal specificity marker *la*. Moreover, in the case of specific indefinites, Mauritian Creole uses both the postnominal *la* and the pronominal *sa*, thus deriving the following construction:

- 71) Pol anvi aste sa zako la.  
 Paul want buy DEM monkey DEF  
 ‘Paul wants to buy this monkey.’

What has always been noted about specificity is the fact that it is nearly tied to (presupposition of) existence. Prince (1981) claims that “specificity necessarily presupposes existence [...], though disputed by Ionin (2006), who argues that specific indefinites assert existence, but do not necessarily presuppose it” (Guillemin 2011:58). Following Guillemin, what results is that a specific definite is familiar to both speaker and hearer and has a strong antecedent in the discourse, whereas a specific indefinite does not belong to the discourse world of the listener and represents new information, as indefinites usually do. Among

specific indefinites, Guillemain introduces an important exception: specific indefinite partitives<sup>60</sup> do have a discourse antecedent, although it is not as strong as in the case of specific definites.

### 3.4 Classification of sentence types

The two types of genericity introduced so far, the nominal and the sentential genericity, can interact at the sentential level, be missing all together, or we may find just one of two, e.g. a kind-referring noun in an episodic (particular) sentence. Recall what we said above about a first difference among generic sentences, namely an internal subdivision in habitual sentences, on the one hand, and lexical characterizing sentences, on the other. As we already said, the former type of sentences generalizes over events and “borrows” verbs from episodic sentential contexts. In contrast, lexical characterizing sentences generalize over properties of individuals. On the basis of the facts described so far, we can now give a first more specific subdivision among generic sentences:

- i. both habitual and lexical characterizing sentences could contain a (non)specific non-kind referring noun phrase or a (non)specific kind referring noun.

As we will see in more detail below, the (non)specificity of a noun phrase reflects the fact that the noun refers to a particular individual or set of individuals (*specific*) in contrast to *nonspecific* NPs which do not have such a reference. A second part of our classification regards episodic sentences, which can be dynamic or statives, depending on what kind of verb they contain. It is important to note here that these types of sentences cannot contain nonspecific predications:

- ii. episodic and dynamic sentences may contain kind-referring or non-kind-referring nouns.

In order to describe other important “generic” facts, we have now to leave aside the sentential level and consider in more detail the reading a kind-referring noun may receive within a sentence. Krifka (1995) makes a list of the various possible interpretations of generic nouns, which we can summarize as follows:

---

<sup>60</sup> Farkas (1994) defines the partitive as a noun phrase, which “denotes a member of a subset of a familiar discourse group” (1994:126 in Guillemain 2011:58).

- 72) a. kind predicate interpretation  
 b. collective property interpretation  
 c. average property interpretation  
 d. characterizing property interpretation  
 e. distinguishing property interpretation  
 f. representative property interpretation  
 g. avant-garde interpretation  
 h. internal comparison interpretation

The interpretation in (72a), the kind predicate one, may only be applied to nouns in conjunction with what we may refer to as *lexical kind predicate*, namely verbs like *be extinct* or *evolve from* and the like. Consider the following sentence from Krifka (1995):

- 73) The dodo is extinct.

In (73) we note the presence of the lexical kind predicate *be extinct*, which can apply to kinds only. All other interpretations listed above are normally derived within sentences containing another type of predicate, namely the *derived kind of predicate*, which can apply to both kinds and objects. Interpretation (72b) refers to the fact that the property described by the verb is true for all the instances specified by the noun phrase. It is thus applied to the whole “collectivity” or group, and not to the kind. The same holds for the *average property interpretation* and can be easily understood by looking at the following example:

- 74) The American family contains 2.3 children.

This sentence is an average estimation of American family units, and is analyzable like the collective property interpretation. Now take a look at the below sentence:

- 75) The potato contains vitamin C.

This is an example of *characterizing property interpretation*, where we find both kinds of genericity. Indeed, the noun phrase *potato* is a kind-referring noun, and the sentence is a characterizing one. The property of the potato described in this sentence, namely the property of containing vitamin C, is characterizing not only for a potato, but for the whole kind as well. This analysis holds for the *distinguishing property interpretation* as well, as it is just a weaker manifestation of what we have claimed above for the characterizing property interpretation. As for the reading (72f), i.e. the *representative object interpretation*, if the sentence includes a

description of a property which is really relevant for the whole kind, it can be applied from the object to the kind. In order to better understand that, consider the following example:

76) In Alaska, we filmed the grizzly.

Here, the grizzly – or grizzlies – filmed in Alaska are representative of the whole species *Ursus arctos horribilis*. In cases where the property under consideration is an episodic property, we are in the presence of the *avant-garde interpretation*, where an exceptional property of some object of a kind is translated to the kind itself, as perfectly exemplified in (77):

77) Man set foot on the Moon in 1969.

Lastly, we have to consider the *internal comparison interpretation*: here no characterizing property of individual specimens is taken into account, but rather “a comparison of the specimens of a kind along a certain dimension of their occurrence” (Krifka 1995:84). Consider the sentence in (78):

78) The wolves are getting bigger as we travel north.

What is changing, namely getting bigger, is not just a couple of wolves, but the whole species *Canis lupus* is undergoing this change in the particular situation described by the predicate, although it is not a characterizing sentence of the kind “wolf”: on the contrary, it is an episodic property linked to the present context under consideration.

### 3.5 Carlson’s (1977) theory

Much work has been devoted to English bare plurals, whose ambiguity between an existential and a universal reading renders it an interesting phenomenon of the grammar. Several studies are concerned with bare plurals, and focus on the ambiguity of their semantics, assuming that such an ambiguity lies in the noun phrase itself. To the contrary, in his treatment of English bare plural nouns, Carlson (1977) makes the claim that the ambiguity of these noun phrases is to be analyzed in terms of a ‘unified phenomenon of the grammar’, and not by means of ambiguity lying in the noun phrase itself. Indeed, he proposes an analysis of bare plurals whose existential or universal interpretations are derived by means of the context of the sentence and the way it acts on the noun phrase. In other words, as explained in

Krifka et al. (1995), Carlson aims at unifying the existential and the generic interpretations of English bare plurals, assuming that the existential one is kind-referring just as well as the generic one.

More particularly, in his work Carlson assumes, on the one hand, that the existential reading is realized by means of the (bare) indefinite plural. On the other hand, we find the universal interpretation, also referred to as ‘generic’. The hypothesis that these two interpretations, the indefinite and the generic, could be due to two semantically distinguished determiners, both taking the syntactic form of  $\emptyset$  (the null determiner)<sup>61</sup>, is left aside by Carlson. In contrast, he assumes that the right way of describing these facts is a unified analysis of bare plurals as ‘names of kinds of things’, which receive the one interpretation or the other by means of contextual factors. Looking at the following sentences can help us better understand that:

- 79) Dogs are good pets.
- 80) Dogs are sitting on my lawn.

As noted in Krifka et al. (1995), both sentences above contain the bare plural noun phrase *dogs*, which has to be analyzed as kind-referring, according to Carlson’s theory. This entails that both bare noun phrases refer to the kind *Canis*. Kind-referring noun phrases can thus also have an indefinite plural reading, as the sentence in (80) shows: it contains indeed a present continuous, a tense that yields an existential reading.

Moreover, Carlson takes notice of an important semantic relationship between bare plurals, or in Carlson’s notation ‘ $\emptyset$ NPs’, and kinds. As a consequence, he suggests a treatment of the bare plural “as denoting a kind of thing”. More specifically, Carlson proposes that “the bare plural acts as the proper name of a kind and that kinds are to be construed as individuals” (1977:442). This assumption is also based on Postal’s (1969) observation that bare plurals and proper names behave the same with respect to the ‘so-called’ construction:

- 81) *Slim* is so-called because of his slender build.
- 82) *Cardinals* are so-called because of their color.

---

<sup>61</sup> It is important to note that Carlson (1977) uses the notion of ‘null determiner’ pretheoretically. He uses ‘ $\emptyset$ NP’ in order to indicate the bare plural noun phrase, without taking position in the discussion about the internal structure of the noun phrase.

Only bare plurals and proper names, indeed, are allowed in such constructions. In contrast, quantified and determined noun phrases are not allowed to occur in such contexts, as Carlson notes. Look at the following examples:

- 83) \*Those/\*All/\*Most/\*No/\*The cardinals are so-called because of their color.

The examples in (83) are not grammatical, as opposed to the sentences in (81) and (82).

As we said above, Carlson (1977) assumes that other elements in the context may influence the derivation of the one or the other interpretation, i.e. the generic plural and the indefinite plural. Among these factors, Carlson crucially notes that tense also plays an important role in the final interpretation of a bare noun. Consider the following examples:

- 84) Dogs are running around in circles.  
85) Dogs run around in circles.  
86) John ate kelp.  
87) John will eat kelp.

Among the sentences above, there are tenses referring to a particular moment of time, and other tenses which can be considered more or less ‘timeless’, in the sense that they describe some habits or features more or less extended in time. The ‘timeless’ reading that speaks of properties, characteristics, habits, etc. does not create any ambiguity, and always selects the universal reading. This is the case of sentences (85), (86) and (87), where we find respectively a present simple, a simple past and a simple future. It is important to notice that the examples in (86) and (87) can also derive an existential reading, as well as the present continuous in (84). Hence, they are ambiguous between a generic and an existential reading. We have already spoken of the former one in terms of ‘timelessness’. On the other hand, as far as it concerns the latter, namely the existential interpretation in terms of the indefinite plural, it will be selected when we find a sentence with reference to a particular time. The simple past and the simple future in the examples (86) and (87) can also be rooted in a specific moment of the past or future respectively, thus describing an event rather than a property. As a consequence, this yields the derivation of the existential reading.

Another important point noticed by Carlson is that tenses are not the only sentential factors that can derive the one or the other reading. This, indeed, can be achieved by means of adjectives as well: some adjectives select the existential reading, whereas some others derive the universal one. Milsark (1979) explains this phenomenon assuming the existence of ‘states’ that select the existential reading (indefinite plural), and of ‘properties’, which, in contrast,

select the universal interpretation (generic plural). ‘States’ are described as more or less temporary, whereas ‘properties’ have a more permanent character. These facts are well illustrated by the following examples:

- 88) Soldiers were available.
- 89) Soldiers are brave.

In the sentences above, we can note that the only interpretation available for the example (88), containing the adjective describing the state of being temporarily ‘available’, is the existential one, which translates into an indefinite plural. On the other hand, the adjective expressing the property of being ‘brave’ yields a generic reading.

Summarizing, Carlson suggests that the predicates which give rise to the existential reading are predicating about an individual at a certain point in the time and space. On the other hand, the predicates which select the generic reading refer to more or less permanent properties (habits, characteristics, etc.) of the individual. At this point Carlson introduces the notion of ‘stage’, defined as the realization of an individual at a certain time and a certain place. Furthermore, “an individual’s set of stages is denoted” as follows (Carlson 1977:448):

- 90)  $\lambda xR(x,j)$

As Carlson explains, the formula above, where ‘j’ stays for the individual ‘Jake’ and ‘R’ indicates a realization relation, has the following reading: “the set of all things,  $x$ , such that  $x$  bears the relation  $R$  to Jake” (ibid.). As a consequence, with respect to our discussion about states and properties, states would be in the property set of a stage of the individual. On the other hand, properties are in the property set of an individual.

Carlson claims that the indefinite plural is associated with an existential quantifier, whose “source is not the determiner of the  $\emptyset$ NP, but rather what is being predicated of it at that time” (1977:451). As Carlson notices, this explains why the existential quantifier will always have the narrowest scope, accounting thus for the fact that it does not interact with other predicates of the sentence. As Krifka et al. (1995) explain, this means that the interpretation is nonspecific, as the following example can clarify:

- 91) Minnie wishes to talk to a young psychiatrist.
- 92) Minnie wishes to talk to young psychiatrists.

The indefinite singular noun phrase in the sentence in (91) can have two different interpretations: a specific one, where there is a certain psychiatrist Minnie wishes to talk to,



and a nonspecific reading, where Minnie has not in mind a specific ‘young psychiatrist’, but rather she wishes there is one and she talks to him. The former interpretation, i.e. the specific one, is achieved by means of a wide scope phenomenon. On the other hand, the latter interpretation, namely the nonspecific one, is due to the fact that the determiner phrase *a young psychiatrist* takes a narrow scope. In the case of the sentence in (92), on the other hand, the situation is quite different and testimonies for the validity of Carlson’s assumption about the fact that English bare plurals always take the narrowest scope, when they receive an existential reading, namely they are nonspecific: indeed, *young psychiatrists* in (92) can only have a narrow scope reading.

In order to demonstrate the validity of his hypothesis, Carlson does not limit his investigation to scopal phenomena, but extends it to anaphora phenomena as well. In this regard, Carlson notes that anaphoric phenomena can take place between existential and generic bare plural noun phrases, unless we meet any ambiguity of interpretation. The same is also true in case the second noun phrase is replaced by a personal pronoun:

- 93) Mick traps *lemmings* even though he knows full well that *they* are protected by law.
- 94) *Lemmings* are protected by law, but Mick goes ahead and traps *them* anyway.

Carlson notes that in the sentence in (93), the bare plural noun phrase *lemmings* receives an existential reading, i.e. it is an indefinite plural and serves as anaphoric antecedent to the generic use of the plural personal pronoun *they*. On the other hand, in (94) we find a generic bare plural referring to the kind *Lemming* acting as antecedent for an indefinite existential use. Anaphoric phenomena also hold in cases where the second element of an anaphoric relation is represented by a reflexive pronoun, as noted by Mats Rooth (1985):

- 95) At the post-WW III peace meeting, Martians presented themselves as almost extinct.
- 96) \*At the post-WW III peace meeting, some Martians presented themselves as almost extinct.

In the sentence in (95), the bare plural *Martians* is in its existential reading, although it refers to the kind *Homo(?) martis*. As explained in Krifka et al. (1995), here the reflexive is co-referential with the subject, and the presence of the kind-level predicate *be extinct* forces it to refer to a kind, i.e. the reflexive is generic. On the other hand, the quantified noun phrase

*some Martians* in (96) is an indefinite noun phrase and therefore cannot refer to a kind; hence, the sentence is ill-formed.

The very same can be predicted of sentences with conjoined predicates, as noted by Schubert and Pelletier (1987). Consider sentences of the following type:

- 97) Snow is white and is falling right now through Alberta.
- 98) Dogs are mammals and are barking right now in front of my window.

As explained in Krifka et al. (1995:116), “one of the conjuncts requires a ‘generic’ interpretation and the other an ‘existential’ one”. Along this line of thinking, *Snow is white* and *Dogs are mammals* receive a generic, kind-referring reading, whereas the second part of both sentences contains a verbal tense, which leads to existential readings only. This would be a piece of evidence in favor of Carlson’s analysis of bare plurals as names of kinds. The different interpretations, i.e. the existential and the generic, would be derived by means of the verbal predicate, without yielding any ambiguity. Furthermore, as explained in Krifka et al. (1995), the existential quantifier is part of the semantics of the verbal predicate and therefore cannot interact with other sentential operators eventually present in the sentence, thus taking a narrow scope. Summarizing, what yields the two different interpretations of bare plurals, the existential and the generic one, is the fact that one conjunct results to be internal to the verbal predicate, whereas the other does not.

### 3.5.1 Some problems faced by the Carlsonian approach

Although there are cases which can be well accounted for by the theory developed by Carlson and described so far, there are at the same time also facts which cannot be explained by means of a Carlsonian view. Krifka et al. (1995) present some of these points, proposing an alternative solution, which considers bare noun phrases ambiguous between an indefinite and a definite, kind-referring interpretation.

A first problem is posited by the fact that there are significant distributional differences between bare plural noun phrases and kind-denoting noun phrases introduced by a definite determiner, both in episodic and characterizing sentences and in *there*-constructions: the two kinds of noun phrases, the bare plural and the definite singular kind-referring, do not always yield the same interpretation. Consider the following examples proposed by Krifka et al. (1995):

- 99) Horses stampeded through the gate.
- 100) The horse stampeded through the gate.
- 101) There are horses stampeding through the gate.
- 102) \*There is the horse stampeding through the gate.

Pairs of sentences like (99) and (100) seem to bring evidence in favor of the alternative solution by Krifka et al. (1995), mentioned above: *horses* in (99) and (101) can be assumed to be an indefinite noun phrase, and not a kind-denoting name as Carlson proposes. Furthermore, the ungrammaticality of sentence (102) seems to be also well accounted for by the fact that the definite noun phrase *the horse* is a kind-referring noun. If grammaticality is to be achieved, the sentence would require an indefinite, non kind-referring noun phrase such as *a horse, horses, some horses, etc.*

Another point which is neither predicted nor accounted for by the theory of Carlson is the fact that bare nouns in their existential interpretation pattern with indefinite noun phrases, thus behaving differently from the bare plurals which receive a generic interpretation.

Consider the following sentences:

- 103) Dogs/A dog are/is sitting on my lawn.
- 104) There are/is dogs/a dog sitting on my lawn.

While the theory of Carlson can account for the case of *there*-constructions by claiming that they are allowed by stage-level predicates, it cannot however predict the similarity of indefinites and bare plurals in their existential interpretation. Even in the case of narrow scope phenomena, the theory of Carlson shows some weakness. Indeed, Krifka et al. (1995) propose an example, whose most plausible reading can be achieved only considering the bare plural under consideration as an indefinite noun phrase which takes wide scope over embedding operators, and which can have anaphoric reference via pronouns in later discourse. Look at the following sentence:

- 105) John intentionally put belladonnas into the fruit salad because he took them for cherries.

As explained in Krifka et al. (1995), the specific bare noun phrase *belladonnas* must take wide scope with respect to *intentionally* and the accusative pronoun *them* must refer to it. The analysis of Carlson would not allow that since, as we said above, the object noun phrase is assumed to be introduced by a verb-internal existential quantifier. The facts are, on the contrary, well accounted for if the bare plural is analyzed as indefinite.

However, there are also arguments in favor of a Carlsonian analysis: it is able to account for conjoined predicates, as analyzed above, and can also explain the lack of ambiguity in sentences with bare noun phrases in subject position, assuming that it is the nature of the verbal predicate that derives the existential or generic interpretation.

### 3.6 Categories and types: On the interpretation of noun phrases

As Partee (1986) explains, there are two major approaches to noun interpretation: on the one hand, we find Montague's description of noun phrases as generalized quantifiers; on the other hand, the second approach distinguishes among referring, predicative and quantificational NPs. Taking as her starting point both approaches described above, Partee sketches a model which connects them and shows that they only apparently contradict each other. Like Montague, she assumes a "requirement of a systematic category-to-type correspondence" (1986:115), but, unlike Montague, she assumes that each category does not correspond to a single type, but rather to a "family of types". According to Partee, this can be formulated as follows: the basic types of noun phrases are  $e$ ,  $\langle e, t \rangle$ , and  $\langle \langle e, t \rangle, t \rangle$ . These types correspond respectively to referential, predicative, and quantificational NPs. As Partee specifies, "while this last, the type of generalized quantifiers, is the most complex, it is also the most general; we can argue that all NP's have meanings of this type, while only some have meanings of types  $e$  and/or  $\langle e, t \rangle$ " (ibid.). Before turning to her treatment of type-shifting operations, Partee takes into account all the possible interpretations of different noun phrases such as proper names, pronouns, quantified nouns, definites, indefinites, and plurals.

As for proper names and singular pronouns, they are described by Partee as being "individual constants and variables respectively" (1986:117). Proper names and pronouns have the basic interpretation of type  $e$ , whereas the generalized quantifier type  $\langle \langle e, t \rangle, t \rangle$  would be achieved via a lifting-rule. This contrasts with Montague's treatment, according to which all noun phrases are assigned the *highest* type, namely that of generalized quantifiers,  $\langle \langle e, t \rangle, t \rangle$ . As a difference, Partee (1986) and Partee and Rooth (1983) propose that "each basic expression is lexically assigned the *simplest* type adequate to capture its meaning" (ibid.). Furthermore, higher type meanings could be assigned by means of certain type-lifting rules, but this happens only if "required in order to combine meanings by available

compositional rules” (ibid.). The next type of NP treated by Partee is the quantified NP such as *every man*. These noun phrases would be assigned only the type  $\langle\langle e, t \rangle, t \rangle$ .

Following Partee, the situation of definite, indefinite and plural NPs is different, since they may have all three possible interpretations,  $e$ ,  $\langle e, t \rangle$ , and  $\langle\langle e, t \rangle, t \rangle$ . In particular, definite noun phrases such as *the man* may have referential, predicative, and quantificational interpretation, respectively:

1. Montague’s Generalized Quantifier interpretation of type  $\langle\langle e, t \rangle, t \rangle$ ;
2.  $\iota$ -operator: it “combines with an open sentence to give an entity-denoting expression, denoting the unique satisfier of that open sentence if there is just one, and failing to denote otherwise” (Partee 1986:117)  $\Rightarrow$  type  $e$ ; and
3. predicative reading: it “picks out the singleton set of (or the property of being) the unique man if there is just one and the empty set (or empty property) otherwise” (ibid.)  $\Rightarrow$  type  $\langle e, t \rangle$ .

As for indefinite NPs such as ‘a man’, they can have all three types of interpretations as in the case of definites, the difference lying in that the predicative interpretation  $\langle e, t \rangle$  corresponds to the bare common noun interpretation. On the other hand, as for the type  $e$ , Partee assumes that indefinites in these cases are “ $e$ -type variables accompanied by conditions on assignments to those variables” (1986:118). In the case at hand, i.e. ‘a man’, the variable must stand for a man.

Finally, as far as plurals are concerned, they may have all three types as well. Partee considers the plural *dogs* and points out its interpretations:

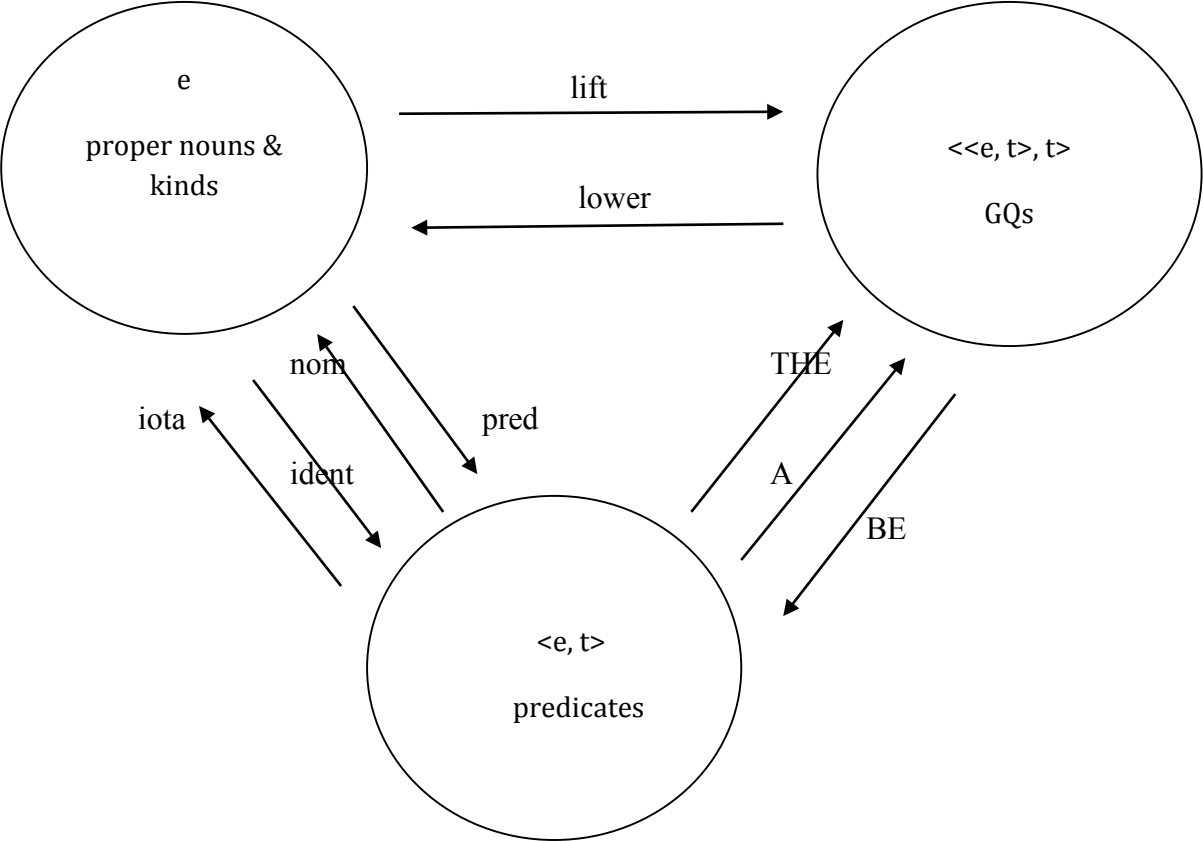
1.  $\cap$  dog’: Chierchia’s nominalization operator  $\cap$  maps “properties onto property-correlates in the domain of entities” (Partee 1986:118);
2. Carlson’s  $\langle\langle e, t \rangle, t \rangle$  interpretation bears the same relation to Chierchia’s nominalization operator (see 1. above):  $\lambda P[P(\cap \text{dog}’)]$  (Partee 1986:116); and
3.  $\langle e, t \rangle$ : plausible interpretation for bare plurals in predicate position.

### 3.6.1 Type-shifting operations: The inventory

In her sketch of type-shifting operators, Partee claims that these are devices which can map “meanings of noun phrases onto other meanings for those same noun phrases”

(1986:121). In her approach to the phenomenon of type shifting, she takes the type  $\langle e, t \rangle$  as the marked type for full noun phrases in English, whereas  $e$  and  $\langle \langle e, t \rangle, t \rangle$  are unmarked types. On the other hand, the type  $\langle e, t \rangle$  is an unmarked option for common noun phrases and verb phrases, and one of the possible types for adjective phrases and prepositional phrases. As already introduced above in the section about Chierchia’s Nominal Mapping Parameter, for an argumental noun to turn into a predicate we need a certain type shifting operation, namely ‘ $\iota$ ’. On the other hand, a predicate turns into an argument when the ‘ $\cap$ ’ operation applies. We have to note here that ‘ $\cup$ ’ and ‘ $\cap$ ’ are not the only type shifting operations available for nouns to change their denotation. Partee outlines, indeed, an inventory of operations, or “type shifting devices that appear to be used in the languages of the world” (Chierchia 1998:358). She does not claim that the set described in her work is a “single uniform and universal set of type-shifting principles” (1986:120), but these operations could be universally available for any language. In more detail, the other type-shifting operations she outlines are as follows (Partee 1986:121):

Figure 1



Partee explains the functions of the type-shifting devices so delineated with the following examples, where ‘j’ refers to an entity, ‘John’:

- i) lift:  $j \rightarrow \lambda P[P(j)]$ , the set of properties that John has;
- ii) lower: maps a principal ultrafilter, if it exists, onto its generator;  $\text{lower}(\lambda P[P(j)])=j$ ;
- iii) ident:  $j \rightarrow \lambda x[x = j]$ , singleton; the set that contains only j as element;
- iv) iota:  $P \rightarrow \iota x[P(x)]$ , provided that P is a singleton set  
iota (ident(j)) = j;
- v) nom:  $P \rightarrow {}^{\cap} P$  (Chierchia); and
- vi) pred:  $x \rightarrow {}^U x$  (Chierchia)  
pred(nom(P)) = P.

In simpler words, we may explain the entire inventory as follows, according to Guillemin (2011):

- Lift turns a proper noun or kind denoting term,  $e$ , into a Generalized Quantifier,  $\langle\langle e, t \rangle, t \rangle$ ;
- Nom ( ${}^{\cap}$  in Chierchia’s notation), whereby a predicate  $\langle e, t \rangle$  is shifted into a proper noun,  $e$ ; it is equivalent to the Nominalizing function  ${}^{\cap}$  treated in Chierchia (1998);
- Pred is the inverse of Nom and corresponds to Chierchia’s Individualizing function  ${}^U$ ; “it assigns existential quantification over instances of the Kind” (Guillemin 2011:86);
- THE is equivalent to the definite article in English, thus turning a predicate  $\langle e, t \rangle$  into a generalized quantifier  $\langle\langle e, t \rangle, t \rangle$ ; this operation is usually assigned the meaning of the determiner *the* itself;
- BE converts a generalized quantifier into a predicate; and
- A is the inverse of BE; it combines English ‘a’ and ‘some’.

Following Partee, these operations are regrouped into three pairs, each one made up of an operation (or more) and its inverse. Moreover, “there is also room for considerable diversity in how natural languages make use of such type-shifting principles, encoding them with

lexical items (*iota* might be a candidate meaning for the definite article), via lexical rules (*nom* and *pred* for the rule relating *blue* as adjective to *blue* as proper noun, depending on which one takes as basic), syntactic rules (*nom* for the formation of bare plurals), or not encoding them at all (e.g. if *lift* is universal for proper nouns)” (1986:123).

Before we turn to the realization of the type-shifting principles described so far, let me introduce another inventory of type-shifting devices, which is slightly different from the set described by Partee. Indeed, the inventory reported in Chierchia (1998:359) is, as the author himself says, an updated version of Partee’s original proposal. Let us begin with the principles which operate a shift from the argumental type  $e$  to the type  $\langle\langle e, t \rangle, t \rangle$  of generalized quantifiers, and vice versa. Here we find Lift and Lower operations, the second being the inverse of the first, as well as in Partee (1986). Lift turns an individual into a generalized quantifier “by taking all of the sets to which it belongs” (Chierchia 1998:359), whereas Lower turns certain GQs,  $\langle\langle e, t \rangle, t \rangle$ , into individuals of the type  $e$ . A first difference regards the shifts from and to the domains of generalized quantifiers and predicates. Here, the device  $\exists$  is an operation of existential closure: its task is that of turning a property into an existential generalized quantifier. This could correspond to the operation A described above. BE is found in Partee’s inventory as well, but in the opposite direction. Following Chierchia, BE is the inverse of  $\exists$ , and, as he claims with Partee, both  $\exists$  and BE are involved in the analysis of copular constructions, or in Guillemin’s words “BE [...] encompasses both the ‘be of identity’ and (predicative) auxiliary ‘be’” (Guillemin 2011:86). Moreover, this type shifting device is traditionally taken as meaning ‘a’, the indefinite article. Finally, we find the operations that turn predicates into individuals and vice versa: ID,  $\iota$ , ‘ $\cup$ ’ and ‘ $\cap$ ’. The latter pair maps properties into individuals and vice versa, as well as the former pair: “the  $\iota$ -operator [...] selects the greatest element from the extension of a predicate and constitutes typically the meaning of the definite article” (Chierchia 1998:359), whereas ID is its inverse part-of relation.

### 3.6.2 Realization of type-shifting operations: Some examples

So far we have reviewed the inventory (or inventories) of type-shifting principles; it is time now to take into account how these devices apply to noun phrases, how they find their realization. Chierchia (1998) assumes that a NP [+arg, +pred] language such as English adopts the pair  $\cup$  and  $\cap$  in order to shift from predicate to argument and vice versa. He explains this choice on the basis of the “Elsewhere Principle”, or “Last Resort”, namely



“Language-particular choices win over universal tendencies”, or in other words “Don’t do covertly what you can do overtly” (1998:360). Following him,  $\cup$  and  $\cap$  are the only pair, which English can use in an “automatic, covert way” (1998:360), since the overt way would be the  $\iota$ -operator for the definite article, and  $\exists$  for the indefinite one, and this overt way would “block”<sup>62</sup> every type shifting operation. This is the case of English and other languages of the NP [+arg, +pred] type, which have articles. Chierchia also considers the case of languages belonging to this group, which, in contrast, do not have articles. This is the case of Russian, where there is no article to block type-shifting operations. Therefore, the  $\cap$  operator has to be used in order to refer to kinds (this is limited, however, to plurals, as well as in the case of English). However, they may also use operators like  $\iota$  and  $\exists$ , as we said above for English: this entails that “bare arguments would occur freely and have a generic, definite, or indefinite meaning, depending, presumably, on the context” (Chierchia 1998:361).

In her treatment of NPs of Mauritian Creole, after having claimed that the noun denotation had shifted from French predicative type  $\langle e, t \rangle$  to MC argumental type  $e$ , Guillemain (2011) takes into account the application of the type-shifting operations as described in Partee (1986). In particular, she describes the realization of the functions lift, nom, pred, THE, BE, and A on MC nouns. Starting with the function lift, Guillemain explains that the realization of lift allows bare mass and count nouns to occur in argument positions and to receive a definite reading. The difference from Partee’s original approach lies in that, according to Guillemain, “the function lift must apply whenever a proper noun functions as an argument, not just in the case of conjunctions [...], when a proper noun like ‘Lilly’ is lifted to the status of GQ so that the conjoined constituents are of the same type” (2011:219). The second function taken into account is the function nom, which shifts nouns from pred to arg: Guillemain argues that it does not apply in MC since nouns have the argumental type as their default denotation. The function pred, in its turn, is performed in MC by the singular indefinite article *enn* and the plural marker *bann*, which derive respectively one instance and multiple instances of a kind. Therefore, pred applies to count nouns only, as specified by Guillemain. Continuing with her treatment of MC noun phrases, Guillemain claims that THE is equivalent to the null definite determiner of MC, whereas the occurrence of the postnominal specificity marker *la*, which marks common count nouns in subject positions thus allowing their definite interpretation, is explained as a ‘last resort’: it licenses the null definite

---

<sup>62</sup> The Blocking Principle intervenes everywhere there results to be a language-particular choice, which therefore wins over universal tendencies (Chierchia 1998:360). In the case at stake here, the eventual presence of a determiner in the language in question would block the application of a type-shifting operation. This last one would be covert, whereas the projection of a determiner is overt, and therefore is preferred to the covert type-shifting device.

determiner in the subject position. Finally, the operation BE is realized in MC by the zero copula, in both copular construction ‘be of identity’ and auxiliary ‘be’, as said above. On the other hand, as opposed to Partee’s (1986) proposal, Guillemín does not consider the function A as a type-shifting device which turns a predicate into a generalized quantifier. Guillemín explains, indeed, that indefinites lack quantificational force and require some other operator, such as tense or a time adverbial, to license their occurrence in argument positions (2011:223).

### 3.7 Chierchia’s Nominal Mapping Parameter

Within the Principles and Parameters framework (Chomsky 1995), Chierchia (1998) studies the characteristics of bare noun phrases from a crosslinguistic perspective, taking into account a revised version of Carlson’s theory of kinds and English bare plurals. Carlson (1977) takes the complex properties of bare plurals as referring to kinds, whereas scholars like Kamp (1981) and Heim (1982), within the framework of Discourse Representation Theory, sketch bare plural noun phrases as being ambiguous between a kind reading and a reading as weak indefinites. With respect to these theories, Chierchia assumes what he calls a “Neocarlsonian view”, whose central point is that bare arguments refer to kinds with no ambiguity between kind referring and weak indefinites. What principally contradistinguishes his approach from Carlson’s (1977) original version is that the Neocarlsonian view does not take into account the notion of stage, defined as a spatial and temporal manifestation of something. Furthermore, with regard to kinds, Chierchia assumes that they are identified by lexical nouns, whereas complex nouns may or may not identify kinds. A way to get from properties to kinds is via the UP ‘ $\cup$ ’ and DOWN ‘ $\cap$ ’ operators, as we will see below.

Chierchia assumes that the denotation of nouns is subject to linguistic variation and “this variation might be responsible for the different distributions of bare nominal arguments” (1998:344). Following Chierchia, the variation lies in the mapping between nominal categories and their meaning: one of the aims of his study is to show how variable this mapping is across languages. Furthermore, Chierchia assumes that nouns have a double function: they can act as predicates or arguments. In the first case, the predicative function, nouns either assume the predicate position or act as restrictors of quantifiers. On the other hand, they are said to be arguments when they refer to kinds. This is so for every language,

the question thus being about the way languages realize these nominal functions, since there is a clear-cut crosslinguistic variation among languages.

Chierchia takes into account different language families in order to classify the way they mark the function of their nouns: Germanic (and English in particular), Romance (taking Italian as example) and Chinese. In explaining how their noun denotation type works, he assumes that  $[\pm \text{arg}]$  and  $[\pm \text{pred}]$  are “features constraining the way in which the syntactic category N (and its phrasal projection NP) are mapped into their interpretations” (1998:353). On the basis of the possible associations among these features, Chierchia classifies languages into three groups: the first one is made up of languages which present the configuration  $[\text{+arg}, \text{-pred}]$ , namely languages whose nouns refer to kinds and, as a consequence, their noun phrases are argumental. Chierchia claims that the languages of this first group allow bare (argument) nouns to occur freely. To the second group there belong languages of the type  $[\text{-arg}, \text{+pred}]$ : here nouns are predicates, and bare arguments are not allowed. Finally, the third group is of the type  $[\text{+arg}, \text{+pred}]$ , where nouns can be freely arguments or predicates. As regards the first group, it contains languages with the following characteristics, such as Chinese: i) they have generalized bare arguments, ii) the extension of all their nouns is mass, and iii) they have a generalized classifier system, but iv) no plural marker. Chierchia claims that these properties are not necessarily logically related and that “in a NP  $[\text{+arg}, \text{-pred}]$  language, all nouns are going to be, in some sense, mass” (1998:353). Furthermore, these languages need classifiers in order to express counting features, as numerals cannot combine directly with nouns. Finally, we have to mention that languages of the type NP  $[\text{+arg}, \text{-pred}]$  are provided with the distinction between mass and count nouns. Indeed, as explained by Chierchia, the fact that all nouns are mass does not imply that these languages do not have any way of distinguishing between mass and count nouns.

With regard to Romance languages like Italian and French, we need to say that they belong to the second group, namely to the NP  $[\text{-arg}, \text{+pred}]$ : this implies that they allow bare nouns in the predicate position only, disallowing bare arguments in most cases, as we will see below. Their nouns are, in fact, predicate of the type  $\langle e, t \rangle$ , following Montague’s notation. One of the distinguishing characteristics of the languages belonging to this group is the fact that the distinction mass vs. count nouns is active, implying thus that some nouns will have a mass extension, and some others a count one. A direct consequence of the presence of count nouns is the availability of plural morphology. The only way these languages have to turn their nouns into arguments is to project a determiner, phonologically realized or not. If D is phonologically non-realized, it is subject to licensing conditions, i.e. governed by a lexical

head. Therefore, if a phonologically non-realized determiner is present, Romance languages allow bare arguments, but only if governed by a lexical head, e.g. the object position. In the subject position bare arguments are disallowed altogether, and a determiner has to be projected.

Finally, we come to the third group of languages described by Chierchia, which includes languages with the following nominal features: [+arg, +pred]. In these languages, nouns freely express themselves as argumental or predicative, i.e. they can denote predicates or kinds, respectively. This is the case for most Germanic languages, e.g. English, which happens thus to be at the middle stage between the other two groups. If an argumental noun has to turn into a predicate, the operation UP <sup>U</sup> applies. Following Chierchia, this would trigger a mass denotation of the noun, which may thus occur as a bare argument, as we have seen for the NP [+arg, -pred] languages. On the other hand, if a noun is predicative, it will be count. As a consequence, the language in question will have a way of overtly marking plurality. Chierchia assumes at this point that a (singular) count noun is not allowed in the argument position if bare, whereas a bare mass noun may occur as an argument.

We come now to the case where a predicative noun has to be shifted into an argumental one: in this case, the shifting operation DOWN <sup>∩</sup> is required. What is being obtained here is a kind, which is, however, possible for plurals only. In the case of a singular noun it is in fact not possible to yield a kind, or in Chierchia's words: "since kinds [...] cannot have a singular instance in every world, <sup>∩</sup> will not be defined for singular properties" (1998:351). So, following Chierchia, if <sup>∩</sup> applies to a singular noun, it will be undefined, whereas if it applies to a plural, it will be defined. In the latter case, the noun is allowed to occur in argument position.

Summarizing, Chierchia assumes that a NP [+arg, +pred] language such as English adopts the pair <sup>U</sup> and <sup>∩</sup> in order to shift from predicate to argument and vice versa. It explains this choice on the basis of the "Elsewhere Principle", or "Last Resort", namely "Language-particular choices win over universal tendencies", or, in other words, "Don't do covertly what you can do overtly" (1998:360). Following him, <sup>U</sup> and <sup>∩</sup> are the only pair, which English can use in an "automatic, covert way" (1998:360), since the overt way would be the *t*-operator for the definite article, and  $\exists$  for the indefinite one, and this overt way would "block" every type shifting operation. This is the case of English and other languages of the NP [+arg, +pred] type, which have articles. Chierchia also considers the case of languages belonging to this group, which, on the contrary, do not have articles. This is the case of Russian, where there is no article to block type shifting operations. Therefore, the <sup>∩</sup> operator has to be used in order to refer to kinds; this is limited however to plurals, exactly as it happens in English. Anyway,

operators like  $\iota$  and  $\exists$  may also be used, as we said above for English: this entails that “bare arguments would occur freely and have a generic, definite, or indefinite meaning, depending, presumably, on the context” (Chierchia 1998:361).

### 3.8 Longobardi’s (1994) proposal

In his treatment of DPs in Romance, Longobardi (1994) moves from the two possible positions for the projection of determiners traditionally taken into account in the literature. These are exemplified as follows:

106)  $[_{NP} DP [_{N'} N]]$

107)  $[_{DP}[_{D'} [_{NP} [_{N'} NP]]]]$

Longobardi assumes the second structure (107) to be the correct one, where the noun is complement of the head D, whereas in the first structure (106) the determiner phrase occupies the specifier position of the noun phrase (Spec, NP). The configuration of this latter structure would not allow the N to D movement, namely the movement from a position inside the NP to a position inside the DP (from specifier position to specifier, or from  $N^0$  to  $D^0$ ): here there would be in fact no c-commanding relation between the two positions involved in the movement. On the other hand, from the structure in (107) it is clear that N can move to D, respecting the necessary licensing conditions. This could explain some particular phenomena of Western Romance and describe them as instances of head-to-head movement. Such a movement would apply, according to Longobardi, in the syntax as far as it concerns Western Romance, whereas it would take place in LF in English and German (1994:609).

#### 3.8.1 A sketch of the Italian nominal system

Italian is a language of the nominal type  $\langle e, t \rangle$ , or in Chierchia’s terms a NP  $[-arg, +pred]$ ; this means that Italian nouns are basically predicates. Hence, for a noun phrase to show up in the argument position, Italian requires the determiner to be projected. This sketch of the situation would cause us to think that a language like Italian, like most Western Romance languages, does not allow bare NPs in the argument position at all. This is, however, not correct. Considering that the subject, direct object, prepositional object, and

inverted subject of either ergative and unergative predicates are the positions where the projection of a determiner is normally required,<sup>63</sup> Longobardi explains that, under certain conditions, Italian also allows BNs in the argument position. More specifically, Longobardi points out three types of bare NPs which can be found in the argument position in Italian, i.e. singular mass nouns, plural count nouns, and (more rarely) singular count nouns in the scope of a sentential negation.<sup>64</sup> The above sketched cases of Italian argumental bare nouns are represented in the sentences below (Longobardi 1994:613):

- 108) Bevo sempre vino.  
‘I always drink wine.’
- 109) Mangio patate.  
‘I eat/am eating potatoes.’
- 110) Non c’era studente in giro.  
‘There wasn’t student around.’

In the examples above, we find a bare mass noun *vino* (108), a bare plural count noun *patate* (109), and a bare singular count noun *studente* in (110) in the scope of the negation *non* respectively. According to Longobardi, bare nouns in Italian seem to be similar to indefinites as far as it concerns their interpretation, and more specifically they are similar to the partitive article, which is made up of the preposition *di* ‘of’ and the definite determiner: the possible combinations are *del/dello*, *della*, *dei/degli*, and *delle*, respectively for masculine and feminine singular and masculine and feminine plural.<sup>65</sup> The most striking similarity lies in that the partitive article can be found with mass nouns and plurals only. There are, however, also important differences, e.g. the behavior of bare nouns with respect to number specification; more specifically, number is irrelevant for mass nouns and negated existential, whereas bare plurals may be neutral between a singular and plural reading (1994:614), whereas a NP introduced by a partitive article always implies a number specification. Moreover, bare nouns “in both Italian and English are subject to an obligatory narrow scope constraint [...] with respect to negation, quantifiers, and intensional contexts” (615).

Longobardi compares bare nouns and empty categories in Italian, since both seem to be subject to a sort of lexical government requirement.

---

<sup>63</sup> As noted in Longobardi (1994:612), Italian does not require the projection of a determiner for noun phrases in nonargument functions, e.g. predicative, vocative and exclamatory contexts.

<sup>64</sup> Longobardi encourages the reader to take notice of the fact that the last case, e.g. singular count nouns in the scope of a sentential negation, is unclear between ‘real argument’ and ‘quasi-idiomatic expression’ (1994:613), as well as in the case of bare nominals as prepositional objects (1994:612 f.).

<sup>65</sup> The partitive articles described so far are syncretic with the prepositional phrases indicating, for example, possession, i.e. ‘of the’.

Generally, bare nouns in Romance are not allowed in the preverbal subject position, but in the internal argument position and, under certain conditions, as inverted subjects of unergative predicates. This is shown in the following examples from Longobardi (1994:616):

111) In questo ufficio incontro sempre marocchini.

in this office I always meet Maroccans

112) Viene giù acqua dalle colline.

comes down water from the hills

113) \*Acqua viene giù dalle colline.

water comes down from the hills

In the sentence in (111) we find a plural count noun *marocchini*, which is allowed in the internal argument position since it is lexically governed. In (112) *acqua* is a bare mass noun, which is allowed in an inverted subject position. With regard to the sentence in (113), it is considered ungrammatical by Longobardi since the bare mass NP is in the preverbal, non-governed subject position; however, it is important to note here that the sentence in (113) is grammatically well formed if *acqua* receives an (intonationally) focused reading.

Furthermore, according to Longobardi's treatment of null determiners as licensing bare nouns, "empty D could instantiate some sort of existential operator and as such impose constraints regarding the count/mass interpretation of the head nouns it quantifies over" (1994: 617). More in particular, he points out that:

- a. Empty determiners are restricted to plural or mass head nouns;
- b. they are subject to lexical government requirement; and
- c. they receive an indefinite interpretation corresponding to an existential quantifier unspecified for number and taking the narrowest possible scope (default existential).

### 3.8.2 The English case

In order to account for English bare nouns, Longobardi (1994) proposes the same as in the Italian case, namely he assumes that a null determiner is present in the nominal structure. English differs from both Romance and the other Germanic languages insofar as the English definite article *the* is neutral between a singular and plural interpretation. That is why some sentences which are not grammatical in Italian can be grammatical in English, as the following examples can show:

114) The secretary and friend of John Smith is/?are coming.

(1994:629)

115) La mia segretaria e tua collaboratrice sta/\*stanno uscendo.

the my secretary and your collaborator is/are going out (1994:619)

Both readings of the English sentence in (114) result to be grammatical at least for some native English speakers, whereas the second reading of the Italian sentence in (115), namely the plural one, is completely ungrammatical, since the definite article introducing *la* is in the singular form and, therefore, only one entity can be referred to. Following Longobardi, the example in (114) shows the presence of an empty determiner like that found in the Italian case examined above. However, it is important to note that the licensing conditions for null D assumed by Longobardi with respect to Italian empty determiners do not have to be considered in the English case. Indeed, English null determiners are not limited to plural or mass nouns, do not seem to be subject to a lexical government requirement like other empty heads, and do not necessarily receive an indefinite interpretation. This is confirmed from a further difference between English and Romance, namely the fact that in English many proper names, which in Romance require an overt determiner, show up without determiner:

116) Old John came in.

117) I love sweet France.

118) Amo \*(la) dolce Francia.

I love (the) sweet France

Another point of divergence between English and Italian also confirms what we said above for English null determiners with respect to the non-validity of the licensing conditions acting in the case of Italian null Ds. Indeed, bare plurals and bare mass Ns in Italian are much more restricted than in English, are necessarily subject to lexical government and receive only an existential reading. In English, instead, bare nouns are syntactically and semantically freer. As a consequence, the existential and generic readings of bare nouns are not always in complementary distribution, although this often happens to be the case. The following example shows that both kinds of reading may be derived with respect to the bare object:

119) I only excluded old ladies.

The adjectival phrase *old ladies* may receive either a generic interpretation or an existential one, depending on the context. In Italian, instead, the situation is quite different because a definite determiner should be used in order to derive a generic reading (or a definite specific).



On the other hand, if we find a bare noun, it will be interpreted as existential, as the sentences below will show. Thus, examples (120) and (121) are respectively the existential and generic reading of the English sentence in (119).

120) Ho escluso solo vecchie signore.

121) Ho escluso solo le vecchie signore.

Moreover, single event verbs like *find* (stage-level predicate) yield an existential reading of the object. On the other hand, individual-level predicates (permanent state verbs) like *love* imply a generic reading of the internal argument. Once again, we have to note that the situation in Italian is different: indeed, the generic reading of bare nouns does not seem to be compatible with Italian individual-stage predicates. A further crucial difference between the two languages under consideration is that English does not require any overt determiner in contexts where, in contrast, Italian does (e.g. with the verb *love*, as we said above).

According to Longobardi, what English and Romance languages share is, however, the fact that both require an overt determiner in the case of generic substantivized adjectives, namely “arguments consisting of an adjective without an overt nominal head” (1994:632), and the fact that singular count nouns must be always introduced by an overt determiner in order to occupy an argument position. The following sentences from Longobardi (1994) contain examples of substantivized adjectives in English and Italian respectively:

122) The rich are becoming even richer.

123) I ricchi stanno diventando sempre piú ricchi.

As a difference from what we said above about singular count nouns, other nominals such as proper names, pronouns, plural and mass nouns can appear in the argument position without an overt determiner. As for the generalization above with respect to the fact that bare singular count nouns are never allowed to show up in the argument position, there are some exceptions. As noted by Longobardi, singular count nouns may occupy an argument position when they receive a mass interpretation; thus, they must be in a condition where their mass reading may be acceptable. More precisely, Longobardi claims that it depends on “their intrinsic meaning and the lexical environment” (1994:633). This could be explained, according to Longobardi, in line with the claim that the empty determiners in Romance and Germanic “seem to impose quantification over subparts and exclude quantification over individuals whenever the head noun following it is in the singular” (1994:633). This property of empty determiners is shared, among others, by the Italian partitive article.

A crucial claim made by Longobardi is that “determiners are semantically understood as operators binding a variable, whose range is always the extension of the natural kind referred to by the head noun” (1994:633). In the case of plural common nouns, such a range consists of members of the extension, whereas as for singulars, the choice between count and mass interpretation depends on the determiner which projects in the structure. From this perspective, Romance and Germanic empty determiners always select the latter of the two options mentioned above, namely the mass interpretation. Plural or mass interpretation of head nouns would be the unmarked (default) option, whereas the singular interpretation, the marked one, would derive from syntactic agreement between the determiner and the noun, whereby a single entity is referred to. Such a ‘singularizing property’ would not be available for empty determiners since it is a marked lexical peculiarity, whereas empty determiners are deprived of lexical content, and therefore DPs introduced by an empty D receive a mass or plural reading, namely the unmarked option. These facts, following Longobardi, would rely on the “denotational interpretation of the DP structure” (1994:634), where the D position is assumed as hosting an operator, whereas the common noun defines a range. Therefore, the fact described above, namely that bare nouns introduced by empty determiners receive mass or plural readings, may directly derive from their kind-referring nature.

Uniquely, proper names and pronouns behave differently, since “they need not refer to a kind and so provide a range to an operator-bound variable” (1994:635): hence, no empty D plays a role in the case of pronouns and proper names. These two types of nominals directly designate the entity they refer to. Pronouns occupy the very same positions at S-Structure in both Romance and Germanic, unlike proper names. Longobardi assumes that pronouns already occupy the D position in D-structure, and therefore no N-raising process happens in the syntax of English.

Longobardi also concluded that nominals are semantically tripartite in pronouns, common nouns and proper names. Such a tripartition is based on the position occupied by the nominals, and on the eventual possibility of movement to the determiner position, i.e. N-raising to D. Indeed, pronouns result to be base-generated in the determiner position, and therefore they can never appear in the N position, since they cannot refer to a kind and therefore they are not able to provide a range to a determiner, be it overt or null, which functions as an operator binding a variable. On the other hand, common nouns are generated in the N position and refer to a kind, thus providing the range described above. This is precisely an inverted situation with respect to that of pronouns. Finally, proper names seem to be at a middle stage between the two nominal types sketched so far: proper names can refer to kinds, and therefore occur in D at least in some languages such as Italian. This means that

proper names can be treated, at least in some cases, as common nouns. According to Longobardi, “in order to refer to a kind [...], a noun must head the N projections at S-Structure” (1994:637), because it is the “irrelevance of the operator-variable interpretation for pronouns and proper names, then, that determines their peculiar distributional possibilities” (ibid.). Furthermore, pronouns are not lexical nouns since they are generated as the spelling out of certain person (and other) features of the head D”. On the contrary, proper names form an open class and “seem to instantiate a more lexical category naturally generated under the N position” (ibid.).

Proper names are semantically different from definite descriptions: as Longobardi claims, “the existential import of proper names seems never to be affected by negation or intensional predicates [...], whereas descriptions instead give rise to frequent scope ambiguities” (1994:638). A peculiarity of proper names is that “an utterance of a sentence containing a genuine referring expression expresses a meaningful proposition only if that expression has a referent” (ibid.). Another peculiarity pointed out by Kripke is the label of rigid designation: proper names seem to designate the same object throughout all possible worlds. Furthermore, Longobardi claims that the D and the N position must be kept as two different entities: thus, they do not belong to the same chain triggering the operator-variable interpretation.

### **3.8.3 Longobardi’s parametric proposal**

At this point, Longobardi makes an important typological generalization, namely that “in languages and constructions where raising of the head noun to the D position substitutes it for the article, only proper names are allowed to raise; in languages and constructions where raising adjoins (prefixes) the noun to the article, common nouns also may be allowed to raise to D” (1994:640). We have, thus, two subcases of movement, i.e. substitution and adjunction, on the basis of which Longobardi develops a parametric proposal in order to explain the different behavior of Italian and English with respect to N-raising: “N raises to D (by substitution) in the syntax in Italian but not in English” (1994:641). This would be an application of Huang’s (1982) proposal, according to which “some languages perform only in LF the same movement operations that other languages already perform in the Syntax” (1994:641). This entails that N-raising to D happens in English at LF, and not in the syntax as in Italian. Before we continue, it is important to specify that Longobardi points out two

universal principles: the first one consists in the fact that the default reading is the existential one, and it is assigned whenever no other interpretation competes. The second principle states that an empty head must be lexically governed.

According to Longobardi, this LF movement that takes place in English may give rise to more types of head nouns to D than the movement taking place at syntax in Italian.

As a consequence, bare plurals and bare mass nouns may also occupy non-lexically governed positions and receive readings other than the existential one. Consider the following example:

124) Big beavers build dams.

The bare plural NP *beavers* is assumed, in Longobardi's analysis, to raise to the D-position, thus substituting it in LF. Such a substitution would prevent a proper government violation. According to this analysis, an ambiguity between existential and generic reading could in some cases arise, as we can see in the following sentence, which we already analyzed above in (119) for the description of English bare nouns:

125) I only excluded old ladies.

The sentence in (125) may receive two different interpretations: an existential one in the sense of 'some old ladies', and a generic one. A third reading would be, actually, possible, as noted by Longobardi. It would be a definite specific reading, which however does not have here any relevance. The ambiguity of this sentence would derive from the fact that the object noun phrase *ladies* is allowed at LF to raise to the position D before the existential reading is assigned. If the noun really raises, a generic reading will be derived. Otherwise, the default existential reading will be assigned, according to that stated in the universal principles discussed above. Both principles are checked as early as possible: the default existential interpretation, for example, is assigned to DPs at S-Structure or at LF, depending on the movement parameter, and cannot be changed in the course of the derivation. The condition that "an empty head must be lexically governed" is also checked as early as possible. More precisely, the assumptions made so far take place in Italian at S-Structure, but may be delayed until LF in English.

As for substantivized adjectives, English behaves like Italian: for the generic reading, an article must be inserted, thus preventing the D position from being empty at LF, since in this case the default existential reading would be assigned. The bare form of substantivized adjectives seems to be rare and depends on the lexical choice of the adjective. In this case, it receives an existential reading and is subject to lexical government constraint, like bare nouns

in Romance languages. The sentences below contain examples of substantivized adjectives which we already analyzed above in (122) and (123):

126) The rich are becoming even richer.

127) I ricchi stanno diventando sempre piú ricchi.

Both in the English sentence and in the Italian one, indeed, the substantivized adjective is introduced by a definite determiner. An important generalization, now, is that bare nouns outside the verb phrase, i.e. in non-internal argument, can only be generically interpreted. On the other hand, as regards bare nouns in initial positions, i.e. in non-governed position, Kratzer (1988) and Diesing (1988, 1989) have proposed that in English the subject of a stage-level predicate can be reconstructed into a VP-internal position at LF, even though occurring in Spec, IP in S-Structure. This position, probably Spec VP, can satisfy the lexical government conditions for empty categories either by virtue of the head V, a lexical governor, or of the head I. Following Longobardi, empty D is in a properly governed position at the relevant level of representation LF in English; it will not need to be filled by N-raising and will receive a default existential interpretation.

### 3.8.4 Expletive articles

The interpretation of generic noun phrases in English is similar to that of singular proper names; both have the following structure, where *e* indicates a trace:

128) [beavers [big *e*]] build dams

129) [water [fresh *e*]] is often drinkable

130) [John [old *e*]] came in

In Italian, the situation is different, in the sense that generics remain in the position N at S-Structure, whereas many proper names can raise to the position D. According to Longobardi, this suggests that generics have some properties of proper names and some other properties of common nouns. On the one hand, they behave like proper names insofar as they “leave no place in D for an overt empty operator (a lexical determiner or a default existential)” (1994:647). This entails that there is no possibility for a DP-internal quantification, “suggesting some sort of barely referential interpretation” (ibid.). On the other hand, since generic nouns are expressions for kinds, they must occur at the N level. Longobardi claims

that generics are names of kinds, as already proposed by Carlson (1977) for English bare plurals.

A peculiarity of proper names that is not shared by generics is that a proper name, especially if not introduced by a determiner, refers “to an entity intrinsically conceptualized as unique in the domain of discourse” (ibid.). As a consequence, if they show up in the plural form, they behave like common nouns, since they do not refer anymore to a unique entity. This behavior is shared by both English and Romance, with the notorious difference that Romance requires the definite article for the generic reading, since bare nouns can receive an existential reading only. A further theoretical statement pointed out by Longobardi is that in the semantics of natural languages we may find only two types of entities, if we leave aside events and states, namely individual objects and whole kinds, “but no subsets of the extensions of such kinds” (1994:648). To these types there correspond the two basic directly referential expressions, namely singular proper names and some generics. This entails that specific plurals are to be obtained via operator-variable structures.

Moreover, Longobardi assumes that the N position refers to kinds, whereas the D position determines the designation of the determiner phrase, either directly or indirectly, i.e. by hosting an operator of a denotational (operator-variable) structure. As concerns the derivation of the correct interpretation, Longobardi claims that the specific reading of a common noun is obtained when the variable, bound by the operator in the D position, ranges over the extension of the kind referred to by the N position. On the other hand, if a proper name not introduced by a determiner raises to D at some level of representation and leaves the foot of the chain, i.e. the N position, uninterpreted, we will have a specific definite reading of the head noun. Thus, DPs headed by proper names in their singular form, directly designate the individual object the name refers to. Generics also create a chain at LF between D and N, but only N is interpreted, the inverse of what we said above for the specific definite reading; the generics so derived have still a referential interpretation. Let us now move to the quantificational interpretation of generics: Longobardi considers that generic DPs can provide the quantificational range to the special invisible operator *Gen* or by adverbs of quantification, which can quantify over a full DP without heading it, similar to *all*: this sort of quantification takes place only with singular and plural indefinite generics in the subject position of individual non-kind-level predicates.

Summarizing, Longobardi assumes that the two relevant positions, D and N, are related by a CHAIN in Italian and that the overt definite article heading generic DPs is an expletive one. In the same way, there are generics introduced by the definite article in English

as well, if the head noun is in the singular, since the nonmass interpretation cannot be expressed through the empty determiner. We have now to recall that proper names are universally able to achieve their specific definite reading without resorting to the operator-variable structure that requires the D position to have substantive semantic content. Thus, the determiners that introduce proper names in some variety of Italian have to be considered as expletive articles and not as having substantive semantic content. A proper name accompanied by an indefinite article receives a kind-referring reading. In conclusion, we may assume with Longobardi that the expletive function of the article is a consequence of the marked, essentially “last resort” nature of the kind-referring interpretation for proper names.

### **3.8.5 Typological evidence**

Unlike Italian, English never allows a definite article to occur with plural and mass generics or with singular proper names. More precisely, English does not allow expletive occurrences of the article, except for singular nonmass generics. Hence, expletive articles are licensed only as a last resort, i.e. if no synonymous raising derivation is available. It is true for English, but it cannot be extended to Romance.

In many varieties of German and in Scandinavian it is possible to use both plural and mass generics and proper personal names with or without the definite article with the same interpretation. Germanic languages pattern with English with respect to the parameter proposed above, namely that “expletive articles are licensed only as a last resort”. This parameter is thus independent from the other parameter proposed for Romance, namely that “N raises to D (by substitution) in the Syntax in Italian but not in English” (Longobardi 1994:641). Thus, according to Longobardi, we have, on the one hand, languages like English and German, and, on the other hand, Romance with respect to the possibility of resorting to N-raising to D in the Syntax: indeed, such a movement is allowed only in Romance, but not in Germanic. At the same time, however, English and German differ with respect to the possibility of resorting to expletive articles; this option is found in some German varieties, as we have seen above. These facts, following Longobardi, could depend on the fact that the English definite article lacks morphological expression of gender and number, whereas the other Germanic languages and Romance have specialized forms for more or less each combination of gender and number. Hence, Longobardi modifies the parameter proposed above for English as follows, turning it into an UG principle: “the phonetic realization of the

D position is licensed only if it expresses semantic content or grammatical features, or as a last resort” (1994:654).

Longobardi considers the definite determiner of English and Italian as “morphological neutralization of two distinct syntactic entities: an expletive and a substantive, really definite, determiner, the latter functioning as an operator” (1994:655). From this point of view, there could be languages which have different forms for each of both types of articles, namely the expletive and ‘real definite’ one. If we recall that two sorts of expletive articles have been treated so far, namely the one which introduces proper names and the other which shows up with generics, we have to say that the first type “relates an interpretively relevant D position (the head of the CHAIN) to an interpretively irrelevant N position (the foot of the CHAIN). On the other hand, the function of the second type is just that of introducing a semantically relevant N position, which refers to a kind, and which lies in an uninterpreted D. Many varieties of Catalan have two types of definite determiner: one is used with proper names and the other in all other circumstances. The following examples will show these facts:

- 131) el gos  
the dog
- 132) en Pere  
the Peter

If the proper name is in the plural, then it has to resort to the plural form of the definite determiner, since “if a name presupposes the possibility of nonsingular designation, it must always be interpreted quantificationally (i.e. like a common noun), and its article can no longer be expletive but must have semantic content as an operator” (1994:656). Another example is the Frisian dialect of the island of Föhr, which has two different types of non-indefinite articles: one occurs with definite specific nominals in both numbers, whereas the other occurs with proper names and with all types of generic phrases. Take a look at the following sentences from Föhr dialect:

- 133) a Türkäi
- 134) Me a deensken san wi leewen frinjer weesen.  
with the Danes we have always been friends
- 135) Di Hans, wat ik käänd, as´r äi muar.  
the Hans that I knew is no more there

Bare nouns, both mass and plurals also exist in Frisian: they may receive an existential or a generic reading. The situation of this dialect is exactly as described above with respect to



varieties with two expletives: both are morphologically neutralized in the A-form, i.e. the form of the determiner *a*, whereas the substantive definite article takes the distinct D-form, i.e. the article *di*. Only the specific reading is available for this last option.

### 3.9 Depréz's Plural Parameter

We are going to introduce a last approach on BNs, i.e. the model built by Depréz (2005, 2006, 2007). In the present section, we are going to briefly introduce it; we will discuss Depréz's Plural Parameter in more detail in Chapter 6, since it will be the starting point of our theoretical implications. According to Depréz, languages may be distinguished into two groups with regard to the category 'number' and the interpretation of bare nouns: the +Plural languages on the one hand, and the –Plural, on the other hand. The presence or absence of a semantic "counter" (2007:322) determines if a language is +Pl or –Pl: more specifically, this semantic counter introduces a number argument which calls for saturation (Depréz 2005). Moreover, the projection of NumP is mandatory in +Pl languages, but not in –Pl languages. We are going to discuss these facts in more detail below and in Chapter 6.

Depréz (2007) claims that, crosslinguistically, the basic denotation of nouns is the kind denotation. Nouns are thus expressions of type <e>; non-kind interpretations are derived by means of "the compositional combination of this basic denotation with a variety of operators introduced by the functional projections of a given language and ultimately by its morphology" (2007:320). This is a crucial difference from Chierchia's (1998) Nominal Mapping Parameter: as we saw above, in Chierchia's model, noun denotation varies according to the ability of the noun in a specific language to be an argument, a predicate or both. Furthermore, according to Depréz (2006), number provides a 'criterion of individuation' mapping a kind onto its realizations: nouns are not default mass; count and mass are distinct realizations of kind.

According to Depréz, the syntactic structure of (bare) noun phrases is essentially the same in both groups of languages: the Number Phrase, indeed, is always present. The crucial difference between the two sets of languages lies in the projection of the Number phrase: NumP obligatorily projects in +PL languages, but not (necessarily) in –PL ones. As we introduced above, a further requirement present in +PL languages is the presence of a semantic counter in NumP: this semantic operator satisfies the number requirement, so that

the noun can be interpreted as singular or plural without giving rise to any ambiguity. As a crucial difference, –PL languages do not (necessarily) have a semantic counter in their NumP. From this crucial difference between the two macro groups of languages, Depréz correctly predicts that singular bare count nouns in the argument position are allowed in –PL languages only. Indeed, since +PL languages always require the projection of NumP, a singular bare noun would result uninterpretable and, thus, be excluded from the grammar.

As for +PL languages, it is important to note, as remarked in Depréz (2007:321), that NumP must always be projected for both singularity and plurality: this entails that what is at stake here is the countability, which is made sure, at least in +PL languages, by the above introduced semantic counter. This latter is defined by Depréz as “a measure function which is the semantic translation of ‘countability’” (ibid.). In +PL languages, thus, we always have number specification for count nouns: the (bare) noun will be either singular or plural, positing that singular possibly means ‘= 1’ and plural ‘> 1’. This may sound trivial, but there are also ‘singular’ forms which are not real singulars. In fact, they are simply underspecified for number. This is the case of –PL languages, after Depréz’s notation. We will see these facts in more detail below.

So far, we have seen that +PL languages completely lack optionality in the sense that they necessarily project NumP. Therefore, the fundamental structure of noun phrases contains a NumP, which is always projected. This holds for bare nouns as well, both singular or (overt) plural or unmarked for number. In contrast, –PL languages do not necessarily require the projection of NumP since their grammars do not always require the presence of a semantic operator. According to Depréz, in –PL languages we may have both the simple projection of NP and the projection of NumP. In case of projection of a simple NP, the result will be a bare singular. We used here the term bare singular to distinguish it from its overt plural counterpart. It would be, however, more correct to speak of ‘bare nouns underspecified for number’. Following Depréz, bare nonplural nouns in –PL languages are instances of what Corbett (2000) calls ‘general number’: they are, indeed, not real singular forms, but they are rather underspecified for number. This implies that they can yield both singular and plural interpretations. Following Depréz, bare number-neutral NPs in all creoles in –PL languages are instances of the ‘general number’. This happens to be exactly the situation of Guinea-Bissau Creole and in creoles in general. Kihm (2007) also speaks of “unspecified for number” with respect to bare singulars in GBC. The underspecification with respect to number may be expressed as follows:

As a difference, if NumP in –PL languages is projected, the semantic counter is ‘activated’ and needs to be saturated: according to Depréz, this can happen by means of overt plural morphology, numerals and an indefinite determiner. Plural morphology, numerals and the indefinite determiner would, thus, encode such a counter.

Following Depréz’s approach, creoles belong to the –PL set of languages. On the other hand, languages such as English, French and Italian would be +PL: as predicted by Depréz’s parameter, such languages do not freely allow bare nouns in argument position; bare singulars are usually disallowed (they would, indeed, be uninterpretable), whereas bare plurals are allowed under certain restrictions. We are going to discuss Depréz’s approach in more detail in Chapter 6. Before we conclude the present chapter, we will speak of BNs in EP, since it is useful to have an idea of how BNs work in Kriyol’s lexifier language.

### 3.10 Bare nouns in European Portuguese: A general sketch

Before we start with our description of Kriyol bare noun phrases, it could be interesting to look at the situation in its lexifier language, i.e. European Portuguese, with respect to bare nouns. EP patterns in this respect with Romance in general. On the other hand, Kriyol shows many crucial differences from EP and Romance in general, as we will see in Chapter 5. For the time being, it is important to note that bare nouns in Kriyol are much more widespread than in EP. In fact, in EP just as in the other Western Romance languages, bare plural count nouns are restricted in their distribution, and bare singular count nouns are found quite rarely. In opposition, bare mass nouns are not a rare phenomenon, at least in lexically governed positions.

A pivotal role in restricting the occurrence of bare nouns in EP is played by the fact that EP has overt (lexical) determiners, both definite and indefinite, which agree in gender and number with the noun they determine. The definite articles, say the equivalent of English ‘the’, are *o/os* and *a/as*, masculine (singular/plural) and feminine (singular/plural) forms. The indefinite articles are, on the other hand, *um/uns* ‘a/some’ for masculine (singular/plural) and *uma/umas* ‘a/some’ for feminine (singular/plural). As a crucial difference, Kriyol does not have any definite determiner. It has an indefinite determiner, which may inflect for number,

although it seems to happen rarely: *un/uns* ‘a/some’. The indefinite determiner in Kriyol does not occur in every context where indefinite articles in Portuguese are found, namely where indefiniteness is to be achieved. As we will see, (in)definiteness finds realization by means of bare nouns in Kriyol, whereas the indefinite determiner is usually found if the noun it introduces is indefinite nonspecific. We will discuss these facts in more detail later on, but for the time being, we just need to notice that EP uses determiners for expressing (in)definiteness. Furthermore, EP overtly distinguishes between singular and plural count nouns by means of overt plural morphology, i.e. the suffix *-s*. In the present section, we are going to see the distribution and interpretation of bare nouns in EP.

According to Müller and Oliveira (2002) and Oliveira and Silva (2008), EP has a ban on bare singulars in argument positions: more specifically, singular count nouns are disallowed in both subject and object positions, although there seem to be some rare cases where a bare singular count noun may show up.<sup>66</sup> Generally, it holds that only bare plurals may show up in argument positions. We need to add, however, that bare plurals are subject to certain restrictions as well: more specifically, following Müller and Oliveira, they cannot be the subjects of episodic, habitual, individual-level, and kind-level predicates. The following examples from Müller and Oliveira show the ungrammaticality of bare subjects of episodic (136), habitual (137), individual-level (138) and kind-level predicates (139):

- 136) ?/\**Amigos partiram ontem.*  
 Friends left yesterday  
 ‘Friends left yesterday.’
- 137) ?/\**Professores trabalham muito.*  
 Teachers work much  
 ‘Teachers work a lot.’
- 138) \**Elefantes são inteligentes.*  
 Elephants are intelligent  
 ‘Elephants are intelligent.’
- 139) \**Elefantes estão extintos.*  
 Elephants are extinct  
 ‘Elephants are extinct.’

---

<sup>66</sup> Oliveira and Silva (2008) mention journal titles as a special case for the occurrence of bare nouns. This seems to be not peculiar to EP only since many languages use a simplified journalism’s language for being short and incisive. For instance, articles are usually avoided, thus much work for the interpretation is left to information structure.

Following Oliveira and Silva, bare plurals in the subject position are allowed with stative non-phase predicates such as *adorar* ‘to adore’ only if bound by a habitual operator. In such cases, it seems that even a bare singular count noun is accepted in the subject position.<sup>67</sup>

140) Turistas adoram viajar.

Tourists love travelling

141) Turista adora viajar.

Tourist loves travelling

With event predicates, bare plurals in the subject position in episodic contexts are allowed and yield an existential reading. Bare singulars are disallowed altogether.

142) \*Elefante invadiu o campo.

‘Elephant invaded the field.’

143) Elefantes invadiram os campos.

‘Elephants invaded the fields.’

If the context is habitual, only bare plurals are allowed, and may yield both existential and generic readings. The kind reading is to be excluded since EP bare nouns cannot yield kind readings, as we already noted above. Furthermore, as we explained in the section about genericity, following Krifka et al. (1995), the kind reading may be yielded with kind-level predicates only.

According to Müller and Oliveira, bare plurals in the object position, on the other hand, are much more free in their distribution. Yet, they cannot yield a kind reading. They usually derive indefinite readings.

144) A Maria compra livros todos os dias.

The Maria buys books all the days

‘Maria buys books every day.’

145) A Maria comprou livros ontem.

The Maria bought books yesterday

‘Maria bought books yesterday.’

146) A Maria lê livros.

The Maria reads books

‘Maria adores books.’

---

<sup>67</sup> The example in (140) seems to be more Brazilian Portuguese than European Portuguese (Kihm p.c.).

- 147) #Os americanos inventaram transistores.  
 The Americans invented transistors  
 ‘Americans invented transistors.’

In defense of our claim that kind-readings are disallowed for EP bare nouns, it is important to note here that *transistores* in (147) does not receive a kind interpretation, but a taxonomic one.

If bare plurals are modified, or in other words if they are made ‘heavy’ (Müller and Oliveira 2002, Longobardi 2000), they can appear bare in the subject position. Following Müller and Oliveira (2002:10), “Longobardi (2000) argues that ‘heaviness’ has the effect of remedying the absence of government of the pre-verbal subject position”.

- 148) Amigos de Coimbra partiram ontem.  
 Friends from Coimbra left yesterday  
 ‘Friends from Coimbra left yesterday.’

This is however not possible for subjects of kind-level predicates:

- 149) \*?Elefantes de grandes dimensões estão extintos.  
 Elephants of big dimensions are extinct  
 ‘Big elephants are extinct.’

Müller and Oliveira conclude that bare plurals in EP are indefinites. In line with Heim (1982), “this means that bare plurals are predicates whose variable gets bound either by a sentential unselective operator or by existential closure” (Müller and Oliveira 2002).

BNs in EP, both singular or plural, are also found as internal argument of light verb constructions, principally after the verb *ter* ‘have’.

- 150) Facto que leva a CGD a considerar que não [CNB]tem obrigações em relação aos trabalhadores.  
 ‘A fact that leads the CGD to believe that it doesn’t have obligations towards the workers.’  
 (Duarte et al. 2010)

According to Märzhäuser (2013), mass nouns may appear bare, as the following examples show. The bare noun *barba* ‘beard’ in (151) is a mass noun in a PP. On the other hand, *ouro* ‘gold’ in (152) is felicitous only if bare. Neither the indefinite determiner nor a numeral can introduce it:

- 151) Sim, as mulheres com barba só no circo!  
 ‘Yes, women with beard in circus only!’ (2013:293; my translation)
- 152) A Maria comprou \*um ouro/\*dois ouros.  
 The Maria bought one gold / two golds  
 ‘Maria bought a gold/ two golds.’ (Müller and Oliveira)

Before concluding this section on EP bare nouns, we should notice, with respect to count nouns, that there are cases where not only bare plurals, but also bare singulars are perfectly felicitous. According to Märzhäuser (2013), coordinated bare nouns in EP, both singular and plural, are allowed in both the subject and object position. But for singular coordinated bare nouns in the subject position, whose interpretation is usually restricted to definite readings, coordinated bare nouns may yield both definite and indefinite interpretations. These facts are exemplified in (153) and (154) from Märzhäuser (2013:292f.).

- 153) Mãe e filha caíram num silêncio prolongado (...).  
 ‘Mother and daughter fell into a prolonged silence.’
- 154) (...) eu ainda assim tenho a impressão que barba e cabelo fica melhor,  
 do que só cabelo grande!  
 ‘I get the impression that beard and hair looks better than long hair only.’

The first sentence yields a definite reading (which is also anaphoric, since the coordinated bare nouns refer to the protagonists of a tale, as reported by Mühlhäuser), whereas the second sentence receives a nonspecific reading.

According to Märzhäuser, EP allows bare plural nouns with indefinite/generic interpretation:

- 155) (...) antigamente, facas e garfos faziam-se de, de(essa), dessa barra.  
 ‘In the past, (even) knives and forks were made of this metal.’

Summarizing, singular count nouns in EP, in both the subject and object position, must necessarily be introduced by an overt D, but for cases such as light verb constructions and coordinated noun phrases. A further case where a bare singular may show up as the subject is with stative predicates like *adorar* ‘adore’, as we saw in (141). Thus, it seems that EP patterns with Italian and the other Western Romance languages in having a ban on bare singulars in the argument position.

As for bare plurals, they are subject to several restrictions, but nonetheless they are more widespread than their singular counterpart. They are usually allowed in the object position, but cannot yield kind readings. As for the subject position, bare plurals may show up in both habitual and episodic contexts. In the former, they yield existential or generic readings, whereas in the latter the existential reading will be yielded. Bare mass nouns do not seem to be restricted in their distribution.

As we will see in the Chapter 4, bare nouns in Kriyol are much more widespread, principally due to the absence of a definite determiner. Both singular and plural bare nouns occur in Kriyol, in both the subject and object positions. Their distribution does not depend, thus, on the predicate type. The predicate type, however, plays an important role in Kriyol as well, with respect to the interpretation of bare nouns.



## Chapter 4

### The nominal system of Kriyol and its bare nouns

#### 4.1 Bare noun phrases and the nominal system of Kriyol: An overview

The aim of the present chapter is to describe and analyse the nominal and determiner system of Kriyol: we are going to look at NPs and functional items with special focus on BNs. More specifically, we are going to study the distribution and interpretation of BNs in argument and non-argument positions. On the one hand, argument positions are external and internal arguments of verbs. In more detail, external arguments are subjects, whereas internal ones are recipient and patient, which would correspond to dative and accusative, respectively, in languages with ‘case’. On the other hand, among the non-argument positions we are going to treat the cases of prepositional objects, topicalized BNs, left- and right-dislocated BNs, clefted BNs, directional objects, and predicates. Consequently, the factors which trigger BN interpretation are going to be analyzed in detail. As for (in)definiteness and (non)genericity, verbal aspect and predicate types play a crucial role, along with the position where a BN shows up. The data upon which the following analysis is based come from my corpus (Rom/Bologna 2008-2009, Berlin 2012, Lisbon 2013).

The present chapter is both descriptive and analytic in nature. We will first look at the nominal and determiner system of Guinea-Bissau Creole. Afterwards, we will examine the behavior of GBC bare nouns, namely their distribution and interpretation. Finally, we will discuss the possibilities of interpretation of BNs and how they are yielded. We will see that predicate type and aspect play a crucial role in the derivation of the correct interpretation of BNs.

Before we start with our description of the nominal system of GBC, we need to briefly consider the co-existence of more subgrammars in the mind of GBC native speakers. Following Kihm (2007), three subgrammars compete in the speakers’ mind. The first one is the basilectal grammar, which is at the one end of the Kriyol continuum. The second one is a modification of the basilectal grammar. This is what is usually defined as Kriyol. Finally, at the other end of our continuum, there is a non-basilectal variety, which is closer to EP. These

subgrammars are available in the speakers' minds, and they can resort to the one or the other quite freely. Distinguishing between these varieties is outside the scope of the present work. Still, it is important to keep in mind that we are dealing with more subgrammars, and not with a "coherent language", at least as for the nominal domain (2007:147).

#### 4.1.1 Definition of Bare Noun Phrases

Creole languages have often been thought of as "special" languages (Bickerton 1981, Thomason 1997, McWhorter 2005, Kihm 2000 among others) not only on the basis of the particular conditions under which they arose, but also for the morphosyntactic characteristics that make them "simpler" (Kihm 2000) than the other historical-natural languages. One of those characteristics is the lack of inflection and (often) of determiners. As a consequence, bare nouns occur particularly often in creoles. In the creolistics, traditional wisdom has often regarded this nominal feature as being one of the characteristics making creole grammars something particular. Other examples of grammatical peculiarities in creoles are morphosyntactic factors like lack of agreement, both nominal and verbal and reduplication.

More recent approaches show the tendency to treat bare nominals as the practical result of a linguistic parameter, which would determine the variation that we can observe between languages that allow bare nouns to occur in argumental positions, e.g. Russian, on the one hand, and languages that necessarily need a determiner in order to do that, e.g. Italian, on the other hand. Russian does not have any overt (in)definite determiner. As a consequence, bare nouns in Russian freely occur in both subject and object positions. On the other hand, Italian has overt determiners, both definite and indefinite. This prevents the free occurrence of bare nouns in the argument position. As we saw in Chapter 3, Italian also allows bare nouns in the argument position, but under several restrictions. As assumed in Chierchia (1998), such crosslinguistic variation would occur in the mapping between nominal categories and their meaning. As a consequence, some languages will have determiners as overt lexical heads of their inner structures, some other languages will not. In order to give a definition of bare nominals, I selected some examples from different creole languages, namely Papiamentu (Spanish/Portuguese based) in (156), Haitian Creole (French-based) in (157), and Sranan (English-based) in (158):<sup>68</sup>

---

<sup>68</sup> Papiamentu is spoken in the Caribbean ABC islands, i.e. Aruba, Bonaire and Curaçao (Kouwenberg 2013). Haitian Creole is, together with French, the official language in Haiti, the Caribbean Sea (Fattier

- 156) Tin **kòmputer** riba mi mesa.  
 Have computer on my desk  
 ‘There is (a) computer(s) on my desk.’ (Kester & Schmitt 2007)
- 157) **Elefan** ap vin ra.  
 Elephant PROG come rare  
 ‘Elephants are/The elephant is becoming rare.’ (Depréz 2007)
- 158) **Ijskasi** no ben de a ten dati.  
 Fridge NEG PAST be DEF time that  
 ‘(A/the) fridge(s) didn’t exist at the time.’ (Bruyn 1995)

Each one of the sentences in (156-158) contains a singular BN. We may easily note that the interpretation of bare noun phrases is not an easy matter and undergoes a more or less high degree of crosslinguistic variation.

Bare noun phrases, however, also show up in many noncreole languages. In (159-161) below, bare (singular) noun phrases occur in argument positions in sentences from Mandarin Chinese, Vietnamese, and Gungbe (Gbe language spoken in Nigeria), respectively:

- 159) Hufei mai **shu** qu le.  
 Hufei buy book go SFP<sup>69</sup>  
 ‘Hufei went to buy a book/books.’ (Cheng and Sybesma 1999)
- 160) Tôi mua **sách**.  
 Isg buy book  
 ‘I buy/bought (a/the) book/books.’ (Kirby 2006)
- 161) **Àsé** jè càzù mè.  
 cat fall pot in  
 ‘A cat fell in a pot.’ (Aboh 2010)

Roughly speaking, we may note that in both creoles and noncreoles, singular BNs have a wide range of interpretation as for the categories number, definiteness, and specificity. In greater detail, they may yield singular and plural, (in)definite and (non)specific readings. Sometimes, genericity is also involved. We are going to see these facts in more detail in Chapter 5. For the time being, it is important to note that, crosslinguistically, it makes more sense to speak of bare singulars, or rather of bare nonplurals, as “real” bare nouns. This topic will be

---

2013), and Sranan is an English-based creole spoken in Suriname, South America. Alternative names for Sranan are Sranan Tongo, Taki-Taki, and Suriname Creole English (Winford and Plag 2013).

<sup>69</sup> SFP stands here for “sentence-final particle”. See Cheng & Sybesma (1999).

exhaustively dealt with in Chapter 6. However, we mentioned that in 3.9, where we described Depréz's (2007) hypothesis that real BNs are unspecified for number.

Traditionally, bare nouns are defined as 'determinerless noun phrases'. As a consequence, they may be both singular and plural. This is the case of Romance and Germanic languages, which however allow bare plurals than bare singulars more easily. As for typologically different languages such as creoles, Kwa and South-East Asian languages, bare nouns may need a more restrictive definition. We are going to use Kihm's (2007:145) definition of bare nouns in Guinea-Bissau Creole as "nouns appearing in their root form, i.e. not marked for number". As we will see in Chapter 5, we can extend this definition to the other languages involved in the present study: Cape Verdean Creole, Santome, Papiamentu, Mandarin Chinese, Vietnamese, and Gbe languages. The case of Brazilian Portuguese seems, on the other hand, to be slightly different. Nonetheless, in order to avoid ambiguities, we are going to use the traditional terminology and distinguish between bare singulars, or rather bare nouns unspecified for number (–[number]), and bare plurals.

## 4.2 Kriyol nominal system

In the following sections, we are going to look at the Kriyol nominal and determiner system. More specifically, we are going to describe GBC NPs, their lack of inflectional specification for gender and their behavior as for number. Certain conditions will turn out to be crucial for the overt mark of plurality, i.e. animacy and referentiality along with discourse relevance (Kihm 1994, 2007). Furthermore, we are going to briefly discuss the realization of count and mass Ns in Kriyol since this distinction entails a certain difference in the semantics of GBC nominals. We will exhaustively discuss this in Chapter 6.

As for the determiners, the Kriyol system is somewhat easier than that of EP. Kriyol has no definite determiner, and may resort to the distal demonstrative *ki(l)/ke(l)* in certain cases. It has an indefinite determiner, i.e. *un*, and the demonstratives *ki(l)/ke(l)* and *e(s)*, which are distal and proximal demonstrative, respectively. We will see in detail in which cases these items are used, since their occurrence is not always mandatory. Finally, we will look at agreement in Kriyol as for modifiers, numerals, and quantifiers. It will turn out that agreement between the noun and the modifier/numeral/quantifier is not the rule in Kriyol, which is a

crucial difference from its lexifier EP. Table 2 illustrates an overview of Kriyol functional items<sup>70</sup>:

Table 2. *Kriyol functional items*

Category	Item	Gloss
Plural suffix	-(i)s	-(e)s
Indefinite D	un(s), utru	a/an, (some), an(other)
Proximal DEM	e/es, esis	this, these
Distal DEM	ki(l), ke(l), kelis/keles	that, those
Quantity expressions	manga de, algun(s)	many/much, some
Universal quantifiers	tudu, kada	all/every, each/every
Numerals	un, dus, tris, etc.	one, two, three, etc.
Prenominal modifiers	bon, mau/mal, etc.;; nha, no, etc.	good, bad, etc.;; my, our, etc.
Postnominal modifiers	burmedju, pikininu, etc.	red, small, etc.

Before we analyse the case of the NP, we should discuss Kriyol functional items. Most items are prenominal, except for *tudu* ‘all’ which is a floating quantifier. As we will see later, this quantifier may occur in different positions with respect to the noun in GBC sentences.

Moreover, modifiers are usually postnominal, except for possessive adjectives and a small

<sup>70</sup> Table 2 is inspired by and adapted from Alleesaib (2012:25-27).

group of quality items (e.g. *bon* ‘good’) which are prenominal. The indefinite D *un* ‘a, an’ has the same form as the numeral ‘one’. *Utru* is described as indefinite D in Intumbo, Inverno & Holm 2013. In my opinion, it is an indefinite pronoun/adjective with the meaning of ‘(an)other’ (162).

- 162) E ta pui (...) **un** galinha na tijela bas, e ta pui mas **utru**  
 3pl HAB put INDEF chicken in bowl below 3pl HAB put more other  
 galinha riba.  
 chicken above  
 ‘They put a chicken on the bottom of the bowl, they put another chicken on the top.’

Numerals, quantifiers and quantity expressions usually occur with singular nouns. The plural suffix *-(i)s* is the only inflectional morpheme. We will look at overt plurality later.

#### 4.2.1 Possible NPs in Kriyol

In the present section, we will take a more detailed look at the Kriyol nominal system and its NPs. We will look at several examples of the possible NPs in GBC: BNs, plural Ns, Ns introduced by the indefinite determiner *un*, by demonstratives, by quantifiers or expressions of quantity, by numerals, and Ns modified by adjectives, both pre- and postnominal. The two sentences below contain BNs (*dinhero* ‘money’ and *kasa* ‘house’), a pluralized N (*djintis* ‘people-PL’), and a noun introduced by the indefinite D, i.e. *un terenu* ‘a land’. These examples show the general tendency of GBC BNs to yield a definite reading in the subject position. The situation for the BN *kasa* in (164) is different: it yields a nonspecific reading exactly as *un terenu* does. Finally, the bare plural *djintis* in (163) yields an existential interpretation.

- 163) **Dinhero** ke **djintis** ta risibi la i pikininu.  
 money REL people-PL HAB receive LOC 3sg small  
 ‘The money that people receive there is not much.’
- 164) Por esemplu pa kumpra **un** terenu, pa kumpu **kasa**, i baratu.  
 for example to buy INDEF land COMP build house 3sg cheap  
 ‘For example, buying a land, building a house, it is cheap.’

As for the DP *e guerra* ‘this war’ in (165), the demonstrative here introduces a noun that has already been mentioned in the discourse.

- 165) N bai pa e guera.  
 1sg go to DEM war  
 ‘I went to the war.’

In sentences (166) and (167) we have expressions of quantity: *tudu* ‘all’ in (166) is a floating quantifier, which introduces a nonplural NP. The same is true for *manga de* ‘much, many’ in (167). It also introduces a nonplural NP. A crucial difference between the two quantity expressions is that *tudu* is a floating quantifier and may occur before or after the N it refers to, and even be separated from it.

- 166) **Tudu** djenti sta abertu.  
 All people stay open  
 ‘All people are/everyone is open.’
- 167) I ten **manga de** kau ku pudi djubi na Guine.  
 3sg have plenty of place REL+2sg can see in Guinea  
 ‘There are many places that you can visit in Guinea.’

As for NumPs, sentence (168) below is no different from the sentences analyzed above in (166) and (167): numerals also introduce nonplural Ns. As we will see later, this is a general tendency in Kriyol, but it is not the rule.

- 168) I ten **dus o tris universidade** na Bissau.  
 3sg have two or three universities in Bissau  
 ‘There are two or three universities in Bissau.’

Finally, we have examples of modifiers: *bon* ‘good’ in (169) and *verdi* ‘green’ in (170) are prenominal and postnominal modifiers, respectively.

- 169) Kila i dja nan **bon** vinhu ki sta la.  
 DEM+LOC 3sg already just good wine REL stay LOC  
 ‘That (there) is already a good wine.’
- 170) Guinea Bissau i tera **verdi**.  
 Guinea-Bissau 3sg land green  
 ‘Guinea-Bissau is a green country.’

### 4.2.2 The Noun Phrase

As a consequence of poor (or rather, almost null) inflectional morphology, both nouns and verbs in GBC are bare, i.e. they appear in their root form. In (171a-d) and (172a-d) we have examples of GBC nouns and verbs, respectively:

- 171) a. *karta* ‘letter’,  
b. *katchur* ‘dog’,  
c. *mindjer* ‘woman’,  
d. *omi* ‘man’.
- 172) a. *skirbi* ‘to write’,  
b. *odja* ‘to see, to look for’,  
c. *laba* ‘to wash’,  
d. *djuda* ‘to help’.

There is no morphological element, which helps us distinguish between the two grammatical categories. The structure in Figure 3, which represents the English noun and verb ‘fish’, may well represent the situation of both nouns and verbs in GBC:

Figure 3                      √ fish

The only inflectional morphology we may find in GBC is the plural marker *(-i)s*: it derives from EP plural morphology *-(e)s*. At this point, it is important to note that, while most creoles tend to be innovative with respect to their lexifier languages with respect to the expression of plurality, GBC has instead taken its plural morphology directly from its superstrate language.

As noted in Kihm (1994), one may face some difficulties in distinguishing between nouns and verbs in GBC. In fact, there is no particular internal element of verbs/nouns indicating whether a word belongs to the nominal category or, instead, to the verbal one. According to Kihm, both phonological and syntactic tests may help distinguish between nouns and verbs in GBC. A first test is based on a phonological feature: verbs are often stressed on the last, open syllable, whereas nouns are stressed on the penultimate when the ultimate is open. There are also (a lot of) cases where the verb is not stressed on the last, open syllable, but on the penultimate, just as with nouns. This is shown in (173), where the verb *bibi* ‘drink’ (with its first singular subject in proclisis) has the same stress as the noun *nbera*:



- 173) *N' bibi* vs. *nbera*  
 ['mbibi] ['mbera]  
 'I drank' '(river) bank' (Kihm 1994)

In such cases, we better look at the syntactic position a word occupies in the sentence. Since GBC is an SVO language, the verb will occur after the subject (pronominal/clitic or lexical) and will be eventually followed by an object. Moreover, the verb may undergo cliticization of both subject and object (Truppi 2009). Whenever the subject is not lexical, a subject clitic (weak pronoun), eventually preceded by a strong pronoun, will precede the verb, i.e. it is in proclisis on the verb. Similarly, objects may also be in enclisis on the verb, which means that nonlexical (clitic) objects, both direct and indirect, can immediately follow the verb.<sup>71</sup> The aspect markers *na* and *ta* are also clitics in GBC,<sup>72</sup> as well as the negation *ka*. They always precede the verb. If both an aspect marker and the negation are required, the negation will precede the aspect marker:

- 174) N ka na odjau.  
 1sg NEG CONT see-2sg  
 'I don't see you.'

Summarizing, the syntactic position of a lexical item helps us distinguish between nouns and verbs. If a word follows an aspect marker or a negation clitic, we are dealing with a verb,<sup>73</sup> but if the word in question is introduced by a quantifier, a demonstrative or modifier such as *bon* 'good', we can easily recognize that we are dealing with a noun. As we will see later, most adjectives, however, are postnominal. As for adjectives, they are less problematic to distinguish. Although they usually do not agree in number and gender with the noun they

<sup>71</sup> Whenever both direct and indirect objects are pronominal, the basic clitic postverbal order in GBC is as follows: V+ indirect object + direct object.

*I na dan el.*

3sg CONT give-1s 3sg

'S/he gives it to me' (Truppi 2009).

According to Kihm (1994), the order indirect object + direct object seems to be the usual order in GBC also in case of lexical objects:

*Mininu manda si mame un karta.*

Boy send his mother a letter

'The boy sent his mother a letter' (Kihm 1994:54).

<sup>72</sup> As we stated in Chapter 2, both *na* and *ta* express imperfectivity: *na* expresses continuity and is used both in the case of progressivity and of the (near) future. On the other hand, *ta* lexicalizes habituality (Kihm 1994:86-98).

<sup>73</sup> As we will see later in the present chapter, there seems to be a class of verbal adjectives in GBC, i.e. adjectives which behave like verbs. More specifically, they may co-occur with aspect and tense markers. Furthermore, whenever a verbal adjective is used in predicative sentences, we do not find any (overt) copula, but rather the verbal adjective will select for a syntactic subject, lexical or clitic. This does not happen with real adjectives, which are introduced by an overt copula and require a lexical subject or a strong pronoun (Kihm 1994:31-41).

modify, they never show up alone. Either they modify a noun or they follow a copula. This is, however, the case of ‘real’ adjectives. As we will see in more detail later, there are also ‘verbal’ adjectives in Kriyol.

#### 4.2.3 The loss of ‘gender’ in the process of creolization

Kriyol does not have sex-based gender distinction, in contrast with what happens in European Portuguese. In fact, EP bases a word’s ‘gender’ on sex. More precisely, the language morphologically expresses gender in both nouns and pronouns. Indeed, EP contains two distinct specialized forms for its third person singular pronoun, namely the masculine *ele* and the feminine *ela*, respectively. The same is also true for the plural counterpart of the third person pronoun. The forms are *eles* and *elas* for masculine and feminine, respectively. As opposed to the situation described above, Kriyol has lost the gender feature and, as a consequence, displays only one form for the third singular person pronoun, i.e. (*el*) *i*, on the one hand, and a unique specialized form for its plural counterpart, namely (*elis*) *e*, on the other.

As far as nouns go, we said above that EP expresses gender via morphological material suffixed to the noun. More precisely, the suffixes *-o* and *-a* are available for the masculine and feminine nominal forms, respectively. Each form also has a plural counterpart, which maintains its gender specification and adds *-(e)s* as plural specification.<sup>74</sup> To better understand the situation, we can look at the following examples: sg.m. *monumento* ‘monument’, pl. *monumentos* vs. sg.f. *pessoa* ‘person’, pl. *pessoas*.

In Kriyol, conversely, we do not find any gender marker. Even the “ending” of a noun does not distinguish between masculine and feminine, but it is part of the stem of the word, as we already noted above in (170) and Figure 2. This means that a noun like *karta* ‘letter’, which directly derives from EP *carta* ‘letter’, has integrated the ending *-a*, which in EP makes a noun feminine. This explains why there is no noun in Kriyol that has a direct gender designation. Yet, GBC grammar allows the overt expression of gender distinctions by means of composition, which is a quite productive phenomenon in this language. As noted in Kihm (1994:126), GBC speakers may create compounds that use *matchu* ‘male’ or *femya* ‘female’ as the second term, such as the following pairs: *ermon matchu* ‘brother’ vs. *ermon femya*

---

<sup>74</sup> The *-e-* in the plural suffix has an epenthetic function. It is used whenever the noun to be pluralized ends with a consonant, e.g. *luz* ‘light’, pl. *luzes* ‘lights’ as opposed to *filme* ‘movie’, pl. *filmes* ‘movies’.

‘sister’, *gatu matchu* ‘tomcat’ vs. *gatu femya* ‘she-cat’ and *dona matchu* ‘grandfather’ vs. *dona femya* ‘grandmother’. As we will see in Chapter 5, other creoles too (e.g. Papiamentu) express sex-based gender by means of composition. For the sake of completion, we have to mention some pairs of related words which, contrary to what has been assumed so far, seem to have a gender specification: *primu* ‘(male) cousin’ vs. *prima* ‘(female) cousin’, and *tiu* ‘uncle’ vs. *tia* ‘aunt’. These nouns derive from their EP counterparts *primo*, *prima*, *tio* and *tia*, respectively. Moreover, it seems that GBC grammar allows derivation by means of EP diminutive suffixes *-inho* and *-inha* (masculine and feminine, respectively): e.g. *kasa* ‘house’ > *kasinha* ‘little house’.<sup>75</sup> However, forms such as *primu*, *prima*, and *kasinha* should not be treated as ‘gender specification’ in its grammatical sense because these nouns do not contain, in fact, a real gender marker. As opposed to the EP case, the endings *-u* and *-a* of Kriyol are taken as a whole from the corresponding EP forms, and not as a combination of root + gender marker.

Before we turn to the determiner system of GBC, we should note that this creole contains noun classes that are based on a gradient of animacy. According to Kihm (2007), such animacy scale is responsible, together with specificity, for the eventual overt pluralisation of nouns in the subgrammar defined as Kriyol. More specifically, nouns with the feature [+human] are always pluralized, whenever more than one entity is implied; [+animate] nouns with high cultural and economical value, such as goats or dogs, also receive the mark of plurality. The other animates pattern together with nonanimates and are usually not overtly pluralized. They are, in fact, more “easily massified” (Kihm 1997:195).

#### 4.2.4 Number requirements

According to Haspelmath et al. (2013), most creoles use overt morphology to distinguish between the singular and plural. There is only a small number of contact languages that completely lack plural morphology (e.g. Korlai and Eskimo Pidgin). Most of the languages that have overt plural markers use them for humans and animates. Finally, a number of creoles are said to behave like European languages and apply invariant plural marking to nouns. Guinea-Bissau Creole is listed among this latter type of language. In my

---

<sup>75</sup> This kind of derivation by means of EP-derived diminutive suffixes was noted during a conversation with a Kriyol native speaker during Tjerk Hagemeijer’s class at the Universidade de Lisboa, where I was a guest PhD student in the spring of 2013. I would like to thank Tjerk Hagemeijer for having encouraged the discussion and Manfred Krifka for further advice.

opinion, this does not fairly describe the situation of GBC, which is much more complex and extremely difficult to count in one group or another. According to Kihm (1994:131-135), plural marking in Kriyol may follow two patterns: the first one is more similar to European languages since it marks nouns as plural whenever more than one entity is referred to. On the other hand, in the second pattern, a noun is overtly pluralized when it is specific and/or animate and if no other lexical item (e.g. numerals or quantifiers) co-occurs to express plurality. It is important to note that these are just tendencies, and, although quite strong, they undergo a high degree of variation. This variation depends not only on the region where the speaker originates, but is also ‘mind-internal’ since every speaker “may resort to one or the other [plural pattern] at any juncture in their current speech” (1994:134).

According to Kihm (2007), a gradient of animacy plays a pivotal role in GBC as to overt plurality. This is, at least, the situation of the variety described in Kihm (1994, 2007). Furthermore, this seems to be exactly the variety spoken by the informants whom I interviewed in Berlin, although it is important to note that they quite freely switch from one subgrammar to another.<sup>76</sup> On the other hand, as for the variety spoken by the Guinean community in Lisbon, the situation is more similar to that described in Haspelmath et al. (2013), probably due to the daily contact with Portuguese. Another fact to take into account is the provenience of the informants: it seems that those who come from Bissau (and from central and southern regions in general) use more overt plural morphology (although the situation is not completely invariant). This is consonant with what Kihm (1994) noted about the first pattern of pluralization. In other words, their use of plural markers is more similar to the situation in EP and in other European languages. Still, speakers from the north, especially from the region around Cacheu, seem to follow the pattern described in Kihm (2007). Plural markers are more likely to occur in [+human] and [+animate] nouns. We are going to see these facts in more detail below.

Before we conclude this preliminary discussion on plurality in GBC, we should mention another important fact: sometimes ‘nonrelevance’ of the (plural) entities referred to within the discourse may override the animacy requirement. According to Kihm (1994), whenever the entities are not conceived of as ‘individuated’ and, therefore, can be seen as sets or groups, they are not overtly pluralized. We assume with Kihm (2007) that the following animacy gradient underlies (at least one) Kriyol subgrammar. Humans and animals with

---

<sup>76</sup> During my M.A. thesis at the University of Florence (Italy), I conducted session interviews with GBC native speakers living in Rome and Bologna. They also seem to speak the variety which Kihm (2007) defines as “Kriyol”. They may, however, switch to one or the other variety.

higher cultural value are easier to individualize, whereas animates such as fish or insects and inanimates are more likely to be conceived of as undifferentiated mass. Yet, inanimates and non-easily pluralizable animates may also be overtly pluralized. In this respect, referentiality plays a crucial role. Roughly speaking, whenever the speaker considers it important to make clear that more than one token (of the entity) is being referred to, the noun will be preferentially pluralized. As we said above, bare (singular) nouns in GBC are unspecified for number. We will see in Chapter 6 how a bare noun takes the singular (semantic) specification or the plural (semantic and/or syntactic) one. For the present discussion, it is important to note that an overt plural in GBC is conceived of as necessarily more than one. Kind-referring nouns are usually not overtly pluralized; referential (object-referring) nouns are more likely to be pluralized whenever plurality (<1) is relevant. If not, we will have a bare singular noun, and if the noun is specific, it is even more likely to be pluralized.

#### 4.2.5 Mass and count Ns

Guinea-Bissau Creole has the distinction mass vs. count nouns. Morphologically speaking, a first opposition between count and mass nouns is that the count Ns may be overtly pluralized, whereas the mass Ns cannot be.<sup>77</sup> According to Depréz (2007), this distinction is reflected in the internal structure of noun phrases. Every noun projects a Number Phrase, which contains a semantic counter, in the case of count nouns. The presence of the semantic counter and its eventual ‘activation’ has important consequences, which we are going to see in detail and discuss in Chapter 6 of the present work. For the time being, it is sufficient to note that mass nouns have no semantic counter.

In GBC, bare mass nouns freely occur in every syntactic position, which is not particular of Kriyol only. Rather it seems to be a quite common phenomenon in natural languages. More specifically, mass nouns crosslinguistically seem to be more likely to occur bare than count nouns. For example, in Romance and Germanic languages, bare mass nouns have a wider distribution than bare count nouns. As we discussed above, in these languages, bare count nouns have a more restricted distribution than in creoles. For example, GBC allows bare count nouns in every syntactic position. As we will see in more detail below, bare

---

<sup>77</sup> This seems to be true for natural languages in general: mass nouns do not denote atomic entities, but rather ‘quantities’ of non-individualizable entities. Mass nouns are not singulars since they do not correspond to one, and cannot be plural since they cannot be more than one. In some cases, we may have plural forms of mass nouns. Forms such as *waters* would, however, yield a taxonomic reading.

singular count nouns in Kriyol are not specified for number. Hence, they may yield both singular and plural readings.

Both mass and count BNs are unspecified for semantic number, i.e.  $\neg$ [number]. This resembles Depréz's (2007) description of the situations for French-based creoles and Cape Verdean Creole. We are going to see these facts in more detail in Chapter 6. Both *masan* 'apple', *arus* 'rice' (175) and *storia* (176) are  $\neg$ [number] and yield a nonspecific reading.

175) Bu ten ke bin ba ta kumpra **masan**, **arus** por esemplu.  
 2sg have to POST BA+TA<sup>78</sup> buy apple rice for example  
 'Then you will have to buy apples, rice for example.'

176) I ka muito tene abitu de kontanu **storia**.  
 3sg NEG much have habit of tell-1pl story  
 'She does not really have the habit of telling us stories.'

As for mass nouns, a point of departure for describing them is to differentiate them from count Ns on the basis of the category number. In fact, mass Ns cannot be pluralized. If they get pluralized, a taxonomic reading is to be derived. Moreover, mass nouns entail atomicity/granularity (Link 1987, Krifka 1989). The mass BN *yagu* 'water' in (177) yields an indefinite nonspecific reading. The BN *yagu* in the object position is particularly suitable for the indefinite interpretation. Moreover, it is the first-mention object, and thus represents new information.

177) E ta tisi bos **yagu** mas na un tijela.  
 3pl HAB bring 2pl.OBL water more in INDEF bowl  
 'They bring you some more water in a bowl.'

Before we discuss Kriyol bare nouns in more detail, we need to look at certain important aspects of GBC nominal and determiner system such as plurality, the indefinite determiner, quantifiers, and modifiers.

#### 4.2.6 Plurality in Kriyol: A tendency

Overt plural morphology exists in Kriyol: the plural marker  $-(i)s$  is the only inflectional morphology present in GBC and, as we noticed above, it derives from EP plural suffix  $-(e)s$ . We already noticed that plurality in Kriyol is a quite intricate matter and, at least

---

<sup>78</sup> The status of *ba ta* is not certain. Whenever in conjunction, the two clitics give emphasis to the discourse.

in one of its subgrammars, the presence of overt plural morphology seems to depend on the relevance of the expression of plurality in the discourse. Moreover, there are two further conditions for overt pluralization: animacy and referentiality (Kihm 2007a), as already introduced above. As to referentiality, Kihm describes referential NPs as denoting “one token or a (potentially) countable set of tokens of a given type” (2007a:150). In this section, we are going to see i) how animacy interacts with the Kriyol plural system, and ii) the role of referentiality. It is, however, important to keep in mind that this is the general tendency in one of the GBC subgrammars. Kihm (1994, 2007a) notes that GBC speakers may resort to more subgrammars and shift from one to the other within the same discourse. There may be several factors triggering the shift among the different varieties with respect to plurality, and we will assume here with Kihm (1994) that it principally depends on discourse relevance and referentiality. It seems that whenever it is relevant with respect to the discourse purpose to overtly express that more than one token (of an entity) is being referred to, the noun will be overtly pluralized. In such cases, discourse relevance and referentiality overrule the animacy requirement. On the other hand, whenever the identifiability of the referent is not relevant, it will be preferentially left unmarked as for plurality. The result will be a bare noun. We will see these facts in more detail below.

We turn now to our discussion about plurality in Kriyol. Depending on the ‘ending’ of the noun to be pluralized, we find *-is* or simply *-s* suffixed to the noun. More specifically, whenever a noun ends in a vowel (178a), the suffix *-s* adds to the noun. On the other hand, when the noun ends in a consonant (178b), the epenthetic vocal *-i-* is inserted between the noun and the *-s* of the plural, like the following comparison can show:

- 178) a. sg. *asasinu* > pl. *asasims*  
 b. sg. *katchur* > pl. *katchuris*

Now we come to the conditions regulating overt plurality in GBC, i.e. animacy and referentiality. As to the former, following Kihm (2007:155), “there seems to be a gradient such that humans are the least abstractable individuals, individualizable animals<sup>79</sup> are more abstractable, and inanimates (including plants) are easily massified”.

---

<sup>79</sup> Individualizable animals are, in the words of Kihm (2007), those animals that have an economical and/or cultural value such that they are likely to be considered the same way as people. Within such categorization we can include animals like dogs, goats, monkeys, etc.

- 179) **Katchuris** pega ku nutru.  
 dog-pl fight with one-other  
 ‘Dogs are fighting with each other.’
- 180) **Garandis** de tabanka ta pega kil, e ta tokal.  
 Old-PL of village HAB take DEM 3pl HAB play-3sg  
 ‘The old men of the village take it and play it.’

In sentences (179) and (180) we find the bare plurals *katchuris* ‘dogs’ and *garandis* ‘old (men)’, which derive from the noun *katchur* ‘dog’ and the nominalized adjective *garandi* ‘big, adult’. *Ku nutru* lit. ‘with another’ is the reciprocal in Kriyol. The above examples seem to be allowed by the fact that we are dealing with “tokens of entity” (Kihm 1994:133) that cannot be seen as an undifferentiated mass. Furthermore, following the gradient of animacy as uttered by Kihm, *katchuris* in (179) are individualizable animals. In other words, they are not easy to group together, or “massify”, because of their cultural value. When we deal with a group or mass, no plural marker is indeed required. Finally, *garandis* in (180) is a [+human] N, which is very likely to receive the mark of plurality.

As for referentiality, the following examples will clarify the situation. The plural noun (*si*) *panus* ‘(her) clothes’ in (181) is referential. On the other hand, *ropa* ‘cloth’ in (182) is nonreferential.

- 181) Maria na laba **si panus**.  
 Mary CONT wash her cloth-PL  
 ‘Mary is washing her clothes.’
- 182) **Ropa** ta labadu ku makina di laba ropa.  
 cloth HAB wash-ed with machine of wash cloth  
 ‘Clothes are (usually) washed with the washing machine.’

In AdjPs, either the modifier is marked for plurality or the noun or both. Crucially, agreement is optional, and referentiality is said to play an important role. Kihm (1994, 2007a) claims that overt pluralization happens on the head N, if referential. However, this does not seem to always be the case. Whereas sentence (183) seems to confirm Kihm’s claim, the example in (184) contrasts with it. More specifically, in (183) the modifier *mindjor* ‘better, best’ is pluralized, i.e. *mindjoris*, whereas the noun *alunu* ‘scholar’ does not receive the mark of plurality. The AdjP *mindjoris alunu* is nonreferential. On the other hand, the noun in (184) is overtly pluralized, and not the modifier: *storias engrasadu* ‘funny stories’ is nonreferential, although the noun is overtly marker ad plural.



183) Antao na ki base e ta pudidu **mindjoris alunu** di jona libertadu.  
 thus in DEM base 3pl HAB put-ed best-PL scholar of area free-ed  
 ‘So, on those bases the best pupils from the freed area are put (in the boarding school).’

184) E ta konta **storias engrasadu**.  
 3pl HAB tell story-PL funny  
 ‘They tell funny stories.’

Moreover, sometimes it happens that AdjPs in which only the modifier is marked as plural yields a referential expression, such as in (185).

185) Kela i **utrus prublema** de pais.  
 DEM+LOC 3sg other-PL problem of country  
 ‘This is another problem of the country.’

We need to take into account a last type of plural pattern with adjectives, namely when both noun and adjective are marked as plural:

186) Li e ta bindi **karus usadus**.  
 LOC 3pl HAB sell car-PL used-PL  
 ‘There they sell second-hand cars.’

In sum, adjectives in Kriyol do not usually agree in number with the nouns they modify: “only the head noun is normally marked for plurality, and [...] only sporadically or in somehow ‘decreolized’ varieties does one run into examples like” (183) and (185) above (Kihm 1994:132; emphasis in original). Thus, the ‘most creole’ case would be sentence (184). Yet, it is important to recall that there are more subgrammars in GBC (or “alternate grammars” as in Kihm 1994:132). Following Kihm, adjectives in the attributive position may agree with the noun they modify, as in (186). In a similar fashion, following Kihm, adjectives in the predicative position may agree with the subject of the predicative sentence. In my opinion, adjectives in the predicative position are preferentially not overtly pluralized.

Normally, when the noun is modified by a numeral, a quantifier or an adjective, it does not receive the mark of plurality. This means that the plural marker usually adjoins only to bare nouns and to noun phrases introduced by a demonstrative, as the following examples show:

187) E ta kushina, (e) ta pui **dus galinha** na tijela.  
 3pl HAB cook, (3pl) HAB put two chicken in bowl  
 ‘They cook, they put two chickens in the bowl.’

- 188) Kil omi ten **manga di mindjer**.  
 DEM man have plenty of woman  
 ‘That man has a lot of women.’
- 189) Durante **ki dias** ainda i ka ta tene nomi.  
 During DEM day-PL still 3sg NEG HAB have name  
 ‘During those days, s/he does not have a name yet.’

In the sentence (187) *dus* ‘two’ is a numeral that modifies the noun phrase *galinha* ‘chicken’. This latter remains in its singular form, although the numeral expresses and implies plurality, in the sense that more than one item is involved. In (188) the quantifier *manga di* ‘plenty of’ does not require pluralization of the noun at its right, i.e. *mindjer* ‘woman’. It is important to notice here that the noun *mindjer* is [+human] and, whenever its plurality is felt as relevant in the discourse, it will be pluralized. Although [+human, +animate] are more likely to be overtly pluralized, this may also happen in the case of [–human, –animate] nouns. Finally, in (189) the distal demonstrative *ki* introduces the plural noun phrase *dias* ‘days’; the mark of plurality in this case is on the noun only. This is, thus, the situation in the subgrammar called “Kriyol”. Again, in more decreolized varieties we may find pluralization on both elements of modified phrases.

Before we conclude the present section, we should note that GBC follows the plural pattern of its lexifier language, i.e. EP. In this respect, according to Kihm (1994:132), GBC (and Cape Verdean Creole) do not pattern with most creoles, which found innovative ways to express plurality compared to their lexifier languages. For instance, Mauritian Creole marks its plurality contrast by means of the prenominal marker *bann*, which derives from French *bande* ‘group’ (Guillemin 2011). Mauritian Creole derives, thus, singular-plural pairs like the following ones: sg. *mo* ‘word’, pl. *bann mo* ‘words’ and sg. *rityel* ‘ritual’, pl. *bann rityel* ‘rituals’. Guillemin (2011:188) explains that the plural marker *bann* is unspecified for definiteness. This state of affairs seems to be similar to the GBC case. Indeed, the plural suffix *-(i)s* is specified only for number and not for definiteness. As for the combination plurality+definiteness, Santome is an interesting case. More specifically, Santome is innovative with respect to plural formation when compared to its lexifier EP: the plural marker in Santome is the prenominal *inen*, which is syncretic with the third plural pronoun. We will see these facts in more detail in Chapter 5. For the time being, it is interesting to note that, in contrast to GBC and Mauritian Creole, the Santome plural marker *inen* seems to act as a plural definite determiner, roughly corresponding to the English ‘the’ (Alexandre & Hagemeyer 2007). This means that Santomé *inen* is specified for both number and

definiteness, unlike Kriyol *-(i)s* and Mauritian Creole *bann*, which are specified for number only.

#### 4.2.7 More on animacy

As we mentioned above, there are two crucial requirements, which trigger the presence of overt plural morphology, namely animacy and referentiality. We have already described these two categories above. In the present section we are going to look at them in more detail by means of some clarifying examples.

- 190) N tene **ermoms**.  
1sg have sibling-PL  
'I have siblings.'
- 191) **Balantas ku Fula** ta da ben.  
Balanta-PL with Fula HAB give well  
'Balanta and Fula people get on well.'
- 192) No tene **diferentes etnias**.  
1pl have different-PL ethnic.group-PL  
'We have different ethnic groups.'

The sentences in (190-192) contain a plural noun each: *ermoms* in (190) occupies the direct object position, *Balantas* in (191) is in the subject position, and finally *etnias* in (192) is also a direct object. However, unlike *ermoms* and *mindjeris*, it is accompanied by the adjective *diferentes*, which is provided with the plural marker just as well as the noun it modifies. In sentence (190), the noun *ermoms* 'siblings' is the plural form of *ermon*, which does not have gender specification, as we already discussed above. The noun *ermoms* receives a plural indefinite interpretation since it represents new information. The object position usually introduces new information, although nouns in such positions may also receive definite readings. We are going to see these facts in more detail later.

The coordinated NP in the subject position in (191), i.e. *Balantas ku Fula*, receives a generic interpretation. *Balantas* and *Fula* refer to the entire ethnic groups and are subjects of the individual-level predicate *da ben* 'get on, get along' in a generic contexts. The genericity of the sentence is yielded by the habitual aspect marker *ta*. This interpretation is perfectly in line with the tendency of Kriyol (bare) subjects to be interpreted as definite, and more

generally to the fact that the left periphery of the sentence in Kriyol seems to be provided with a certain degree of definiteness. This is, however, not the case of Kriyol only, but rather it seems to be a quite natural tendency in the information structure of natural languages in general. The definite reading for the subject position would be, thus, the default case, if not otherwise specified, for example from the presence of an overt determiner, be it indefinite or demonstrative. On the other hand, the sentence in (192) contains a plural modified noun phrase that receives an indefinite reading: *diferentes etnias* ‘different ethnic groups’.

So far we have discussed examples of [+human] nouns. Now we are going to take a look at the situation as for nonhuman animates.

193) N tene **santchus**.  
 1sg have monkey-PL  
 ‘I have monkeys.’

194) E ta bindi **pis**.  
 3pl HAB sell fish  
 ‘They sell fishes.’

*Santchus* ‘monkeys’ in (193) seems to behave like [+human] nouns, whereas *pis* ‘fish’ in (194) is more likely to get a mass interpretation. The situation described for *pis* is not unexpected since ‘fish’ is not easily individualizable. Therefore, they occupy a lower position in the animacy scale. It is worth noting that English nouns such as ‘fish’ and ‘sheep’ are usually not marked for plurality, e.g. ‘five fish’ and ‘five sheep’. Whenever we find overt plural forms such as ‘five fishes’, a taxonomic reading is yielded (Manfred Krifka p.c.). In our gradient of animacy, below animals such as ‘fish’, we find plants and inanimates. The following examples demonstrate the behaviour of such nouns as for plurality:

195) Patrik i ta kumpra **planta** na Senegal i leva pa Guiné.  
 Patrik 3sg HAB buy plant in Senegal and bring to Guinea  
 ‘(As for) Patrik, he buys plants in Senegal and brings them to Guinea.’

196) N ka ta toma **mesinhu** pabia n ka gosta de (toma) **mesinhu**.  
 1sg NEG HAB take medicine because 1sg NEG like to (take) medicine  
 ‘I do not take medicine because I do not like (to take) it.’

Both *planta* ‘plant’ (195) and *mesinhu* ‘medicine’ (196) are interpreted as plurals, but they show up in a nonplural form. It depends on the fact that they are easily massified in such cases. Moreover, they are internal objects of predicates accompanied by the habitual marker

*ta*, and habitual contexts are more likely to yield a generic reading of bare nouns than episodic ones (Krifka et al. 1995, Baptista 2007). Before concluding, we should note that the situation described above as for *mesinhu* ‘medicine’ is not typical of Kriyol alone. There are languages such as English which also use a nonplural form in the same context. On the other hand, this does not happen in languages such as Italian. Although ‘medicine’ can be easily massified, Italian requires a plural form.

In conclusion, we may say that there is a strong tendency in GBC to overtly mark nouns as plural whenever the plurality of their referents is relevant to the discourse. Thus, it will depend on the eventual aspect markers present in the sentence. As we saw above, the habitual aspect marker *ta* usually yields nonspecific readings. On the other hand, episodic sentences and the eventual presence of the continuous marker *na* will yield a specific reading.

#### 4.2.8 The absence of the definite determiner

We already said that Kriyol nouns occur bare most of the time or, in other words, they appear in their root form unaccompanied by any determiner or specification for gender and number. One of the factors determining this situation in GBC is the lack of any definite determiner. BNs in Kriyol are typically interpreted as definite, at least in the subject position. Looking at the following examples will provide an idea of the situation:

- 197) **Renda** na karu kada byas mas.  
 rent CONT expensive each time more  
 ‘The rent is becoming more and more expensive.’
- 198) Kuma ke **festa** kuri?  
 how REL party run  
 ‘How was the party?’
- 199) **Badjuda** i bonitu.  
 girl 3sg nice  
 ‘The girl is nice.’

The sentences in (197-199) all contain a bare noun phrase, i.e. *renda*, *festa* and *badjuda*, respectively. They receive a definite interpretation, although they are not headed by an (overt) definite determiner. This is true in the sentences above, and more generally when a bare noun occupies the subject position. As we will see later, there is a high degree of variation in the

interpretation of bare nouns in Kriyol, depending on various factors, such as, first and foremost, their syntactic position in the sentence. We are going to see these facts in more detail later. Here below we find other examples of definite BNs:

- 200) **Dinossauru(s)** ka ten dja.  
 dinosaur(-PL) NEG have already  
 ‘Dinosaurs are extinct.’
- 201) Parsin kuma **salariu** sta in dia, n ka tene sertesa.  
 seem-1sg COMP salary stay in order 1sg NEG have certainty  
 ‘It seems to me that salaries/the salary are/is guaranteed, I am not sure.’
- 202) **Sol** sai.  
 sun rise  
 ‘The sun has risen.’

*Dinossauru* ‘dinosaur’ in (200) yields a kind-referring reading. Generic sentences with a kind-deriving predicate, such as ‘to be extinct’, typically derive such an interpretation. Native speakers agree on the possibility for *dinossauru* to take its plural form without changing the meaning of the sentence. The BN *salariu* ‘salary’ in (201) receives a generic interpretation which is quite typical of generic sentences; however, we already noted in Chapter 3 that, although generic sentences typically yield generic interpretations of their BNs, it may also be the case that the BN receives a nongeneric reading. Finally, *sol* ‘sun’ in (202) receives a definite reading. This is very typical of unique-entity BNs such as ‘sun’.

There are, however, cases where a sort of overt determiner may be found. According to Kihm (1994), in the context of a determinate reference or in case the NP is anaphoric, the distal demonstrative *kil* ‘that’ may precede the noun. Following Kihm, the presence of *kil* in such contexts is not indispensable. It seems to be a simple option of Kriyol grammar, and the choice is left up to the speaker. The examples below will clarify the use of *kil* in determinate contexts, i.e. followed by an attributive relative clause (203) and in anaphoric function (204):

- 203) N kunsá na lei (**kil**) **libru** ku bu piston.  
 1sg start to read (DEM) book REL 2sg lend-1sg  
 ‘I began to read that book that you lent me.’
- 204) Maria pensa kuma **filmi** ku Paulo konsidjal pa i djubi i bonitu.  
 Maria think COMP movie REL Paulo recommend-3sg COMP 3sg see 3sg nice  
 ‘Maria thinks that the movie that Paulo recommended her is good.’

Sentence (203) is perfectly grammatical even without *ki*. On the other hand, in (204) a bare noun shows up and is perfectly felicitous without being introduced by the distal demonstrative. This confirms the optional presence of *ki* in contexts of determinate reference.

A second context where the demonstrative *ki* may act as a definite determiner in Romance or Germanic languages is the anaphoric context. Whenever the referent of the noun is already known in the discourse world of the speaker and hearer, namely if it has been already mentioned, the noun may be introduced by *ki*. The first mention of a referent is usually expressed by means of bare nouns. In this case, although the use of the demonstrative is optional, it seems to be strongly preferred. These facts are clarified in the following sentences:

205) E ta distindi **ki** stera.  
3pl HAB lay.out DEM mat  
'They lay out the mat.'

206) E tisiu **ki** kabas ki tene **ki** galinha.  
3pl bring-2sg DEM bowl REL have DEM chicken  
'They bring you the bowl with the chicken.'

In the case of (205), *stera* 'mat' was previously introduced in the discourse context, and *ki* is perfectly felicitous in its anaphoric role. Its use is, however, not mandatory. The example in (206) may at first sight resemble the one in (203) above. Both display context of determinate reference, but in the second case *ki kabas* 'the/that bowl' is anaphoric. In fact, the bowl was already mentioned in the discourse. In (203), on the other hand, *(ki) libru* is not anaphoric: there was no previous mention of it.

Before we conclude this section, we should mention that there is another case where demonstratives may be used, namely in case of first and subsequent mentions of an object. We will address this in more detail later.

#### 4.2.9 The indefinite determiner *un*

Indefiniteness is normally associated with the concept of 'newness'. As noted in Kihm (1994 :135), "the definite-indefinite contrast as it manifests itself in languages like English or Portuguese is often correlated with the distinction between what is presupposed and what is new in the common world sustained by the participants". More specifically, the use of an

indefinite signals the introduction of a new discourse referent. This typically implies that the object that the discourse referent is anchored to is different from the object mentioned so far. Similarly, indefiniteness is associated with newness also in the case of Kriyol. Nonetheless, we should take notice that, unlike English or European Portuguese, Guinea-Bissau Creole often uses a bare noun phrase where indefiniteness is implied, namely where English and EP would use *a/an* and *um/uma*, respectively. The same is true for the plural case, where English uses either the indefinite plural quantifier *some* (see Carlson 1977) or bare nouns in their existential reading. Similarly, EP uses the plural counterpart(s) of the indefinite determiner, which also agrees in gender with the noun it introduces: *uns/umas* ‘some’, masculine and feminine forms, respectively. In some cases, EP also uses BNs, as we already saw in 3.10. Therefore, indefiniteness in GBC is usually expressed via BNs. In some cases, indefinite nouns are introduced by the indefinite determiner *un*, as we are going to see in the present section. We should recall that there seem to co-exist more varieties of GBC which may differ at some level as to the use of *un* and of overt pluralization. That said, it seems that *un* is unspecified for specificity.

As described in Kihm (1994), in GBC the form *un* and its plural counterpart *uns* ‘some’ are derived from the EP indefinite determiner *um*, *-a*. In EP too we find the same forms for the indefinite determiner and the numeral. As an important distinction from Kriyol, the determiner in EP has to agree in number and gender with the noun it introduces: *um livro* ‘a book’ vs. *uns livros* ‘some books’ for singular and plural masculine, respectively; and *uma luz* ‘a light’ vs. *umas luzes* ‘some lights’ for singular and plural feminine, respectively. As we already stated, Kriyol has no gender distinction. Therefore, *un* and *uns* can be both masculine and feminine, depending exclusively on the context. As a consequence of limited agreement in Kriyol, *uns* is not expected to show up in every plural context. Moreover, its occurrence seems to be quite rare. The most common pattern is the following:

- 207) Kada jona ke (e) padidu i ten **un** **mantenhas** fixu.  
 each area REL (3pl) be.born-ed 3sg have INDEF family.name-PL fixed  
 ‘Depending on the region where people were born, they will have certain fixed surnames.’

As an important difference from both Romance and Germanic languages, GBC seems to have a tendency to overtly mark the contrast specificity vs. nonspecificity more than the contrast newness vs. presupposedness. According to Kihm (1994:135-140), the indefinite determiner *un* is used whenever the noun is nonspecific. It would be, thus, semantically associated with nonspecificity. On the other hand, indefinite BNS are unspecified as for specificity. Following



Kihm (1994:137), it seems, indeed, that “the speaker considers (in an implicit fashion, of course) that the specificity issue is simply not relevant in the [...] context”. The nonspecific reading of *un*-DPs in GBC is exemplified in the following sentences:

- 208) N dibi di kumpral **un prenda**.  
 1sg have COMP buy-3sg INDEF present  
 ‘I have to buy him/her a present.’

In principle, a DP like *un prenda* ‘a present’ in (208) could be ambiguous as to specificity. If we think of Italian, the same sentence as in (208) could yield both a specific and a nonspecific reading:

- 209) Devo comprarle **un regalo**.  
 Must.PRS.1sg buy-her INDEF present  
 ‘I have to buy her a present.’

The DP *un regalo* ‘a present’ in (209) could yield both a specific and a nonspecific reading. The speaker could have in mind a particular object to buy as a present, or could be just saying that s/he should buy a present, but s/he does not know yet what to buy. Now, we come back to our GBC example in (208): the ambiguity of the Italian sentence in (209) is not present in (208) since *un*, as we said above, is inherently specified for nonspecificity. As a consequence, *un prenda* can yield a narrow reading only. On the other hand, in the case of indefinite BNs, scopal properties may play a crucial role in distinguishing between a specific and nonspecific reading. Below, we have two sentences containing an *un*-DP each: *un lesu* ‘a lace’ in (210) is nonspecific. This is consonant with what has been said so far. As a difference, *un riu* ‘a river’ in (211) is a specific DP, but crucially, it is modified by a relative clause. These examples confirm the GBC tendency to interpret *un*-DP as nonspecific, if not modified. It is important to keep in mind, however, that it is just a tendency, and not the rule.

- 210) I ta rapadu kabesa, i ma raka **un lesu** na kabesa.  
 3sg HAB shave-ed head 3sg more tie INDEF lace in head  
 ‘His head is shaved, he also ties a lace around his head.’
- 211) Bissorã tene **un riu ke** rudial.  
 Bissorã have INDEF river REL surround-3sg  
 ‘Bissorã has a river that surrounds it.’

As for new discourse referents, they are usually introduced by BNs. Subsequent mentions of them, both modified (213) and not (212), are usually introduced by *un*. The DP *un panu* ‘a

towel' in (212) is the second mention of the object, which was formerly introduced by a BN. The *un*-DP yields an indefinite nonspecific reading. In the second sentence of (213), *un ropa branku* 'a white dress' is the second mention of the object, and is modified by an adverbial phrase. It receives a specific interpretation. Sentences such as (211) above and (213) lead us to an interesting generalization. When indefinite NPs are modified by a relative clause (211) or by an adverbial expression (213), *un* introduces it. Finally, the first-mention object *ropa branku* in (213) is definite since it is unique referent.

212) U ta maradu **panu**. (U) ta mara **un** panu asim di bas.  
 2sg HAB tie-ed towel 2sg HAB tie INDEF towel so of below  
 'A towel is tied around you. A towel is tied around you below there.'

213) E ta bistiu dja **ropa branku**.  
 3pl HAB wear-2sg already cloth white  
 I ten **un** ropa branku **propri pa ki serimonia**.  
 3sg have INDEF cloth white exactly for DEM ceremony  
 'They dress you up with the white dress. There is a (typical) white dress exactly for those ceremonies.'

Before we conclude this section, we should touch on specific and nonspecific indefiniteness. Kihm (2011) notes that both in Kriyol and in the local Atlantic languages there is a contrast between two kinds of indefiniteness: "the bare truly indefinite or generic nouns ('any x') and noun phrases involving an item that conveys specific indefiniteness ('a certain x')" (2011:90). In Kriyol, this contrast finds realization in the indefinite determiner *un* with the peculiarity that the interpretation 'a certain goat' may also be conveyed by noun phrases headed by *utru*, which has the double meaning of 'a certain' and '(an)other'. Kihm explains that the expression *utru kabra* could mean both 'another goat' and 'a certain goat'. This overlap would be evidence of substrate influence in Kriyol. More specifically, it seems to be typical of languages like Balanta and Manjaku. The sentence in (214) is an example of *utru* with the meaning of 'some, certain'. The meaning 'other' is infelicitous since the sentence belongs to a context where the speaker is describing agriculture as means of support in Guinea-Bissau. She is saying that there are people who almost exclusively live off of what they produce. The indefinite adjective *utru* would have thus an existential meaning. The same is true for (215): *utrus pepeles* cannot yield the meaning 'other Pepels' since the speaker had not mentioned Pepels yet. The speech is about the wedding ceremony's tradition of Christian people in Bissau and focused on Balantas. Sentence (215) is the first mention of Pepels in the discourse context.

- 214) **Utru** djentis ta vivi son di ...di kel ke ta labra del.  
 INDEF people-pl HAB live only of... of DEM REL HAB cultivate of+3sg  
 ‘Certain people live only off of what they cultivate by themselves.’
- 215) **Utrus** pepeles tam gosi ke sta dja na Bissau mismu,  
 INDEF Pepel-pl also now REL COP already in Bissau exactly,  
 ke na vivi dja na Bissau e ka muito ta fasi  
 REL CONT live already in Bissau 3pl NEG very HAB do  
 ki tradison de pepeles la di mas de *aldeia*.  
 DEM traition of Pepel-pl LOC of more of village (Port.)  
 ‘Certain Pepels too, who are already in Bissau, who already live there in  
 Bissau, they do not follow very much that tradition of Pepels from the village.’

#### 4.2.10 Demonstratives

Guinea-Bissau Creole demonstrates overt means for expressing deixis, anaphoric and determinate reference, namely the demonstrative determiners *es* and *kil/kel*, which are the proximal and the distal demonstrative, respectively. Both demonstratives have shorter forms, i.e. *e* and *ki/ke*, respectively. They go back to EP demonstratives: *es* derives from EP proximal demonstrative *este*, *-a* ‘this’ (masculine and feminine), whereas *kil/kel* come from the distal one *aquele*, *-a*, ‘that’ (masculine and feminine). It is important to note that neither the proximal nor the distal demonstrative seem to be allowed to receive any mark of plurality. This is true whenever the demonstratives are used in their function of determiners. The situation is different whenever *es* and *kil* are used as pronouns. In this case, if pluralisation is required, we find the plural forms *esis* ‘these (ones)’ and *kelis* ‘those (ones)’.

So far, we have introduced both uses of GBC demonstratives, namely the adnominal (determiner) and the pronominal function. We already discussed the role of the determiner *kil* in the cases of anaphoric and determinate reference. As for the adnominal function, GBC demonstratives act as determiners with strong deictic force; they are thus strong determiners (Schwarz 2009). On the other hand, in their pronominal function they refer to individuals and substitute for them. As explained in Guillemain (2011:65), “demonstratives serve to focus the hearer’s attention on particular objects”. Depending on how they do that, we can have an *exophoric* or an *endophoric* use of these means for deixis. In fact, if the attention is focused on objects outside the discourse, we have the first type of use and can speak of ‘pointers’. On

the other hand, when demonstratives point to discourse-internal objects, they have an endophoric use. The examples below will show the adnominal use of the proximal demonstrative *es*:

- 216) Mas ke **es** **situason** ki djentis na pasa...**es** **guera** i kusa  
 but with DEM situation REL people-PL CONT pass...DEM war 3sg thing  
 ke manga de djenti ta medi bai investi la.  
 REL plenty of people HAB be.afraid go invest LOC  
 ‘But with this situation which people are going through...this war is why a lot  
 of people are afraid of investing there.’
- 217) **Es** **binhu** ta bibidu friu.  
 DEM wine HAB drink-ed fresh  
 ‘This wine is to be drunk cool.’

In the sentence in (216), we find two cases where *es* occurs as a determiner introducing a noun phrase, i.e. *es situason* ‘this situation’ and *es guera* ‘this war’. The sentence has been uttered once the speaker had already introduced the topic of the difficult situation of the country. The use of the demonstrative(s) is thus endophoric and anaphoric, since it has an antecedent within the discourse. On the other hand, the sentence in (217) contains a proximal demonstrative introducing *binhu* ‘wine’, which does not find explanation in the discourse. Here, *es* should be seen as a ‘pointer’ which focuses the attention outside the discourse. A different case is shown in the sentence below, which will exemplify the case of the pronominal use of the proximal demonstrative:

- 218) N ka kungsi **es** o **es** i de tal tera, ka i...  
 1sg NEG know DEM or DEM 3sg of certain land, NEG 3sg  
 tudu anos i Guineus.  
 all 1pl.TOP 3sg Guinean-PL  
 ‘I don’t know (situations like) this one or this other one comes from that  
 certain land, it is not... we all are Guineans.’

In (218) the two proximal demonstratives are used pronominally, namely they replace the noun. In this case, the noun could have been ‘man’ or ‘guy’ or the like. They are used with a kind of specificity, although the individuals referred to are just a hypothesis in the discourse, and not real entities.

We come now to the case of the distal demonstrative *kil*, which may occur under the same conditions as the proximal *es*. Take a look at the sentences below:

- 219) Na guera na **kil ladu** e kuri tudu.  
 in war in DEM side 3pl run all  
 ‘During the war, in that side (of the county) they all run away.’
- 220) I tene **ki shapeu** ke e ta usa.  
 3sg have DEM hat REL 3pl HAB use  
 ‘There is that hat that they (typically) use.’
- 221) Dan **ki libru**.  
 give-1sg DEM book  
 ‘Give me that book.’

In the sentence in (219) *kil* introduces the noun *ladu* ‘side’ with the function of pointing to a situation which had been already mentioned in the discourse, or at least which is implicit. The speaker explained why elephants are not to be found in Guinea-Bissau: they ran away during the war because they were massively hunted. Hence, *kil ladu* has a determined spatial reference internal to the discourse. As for the sentence in (220), *ki* introduces the noun *shapeu* ‘hat’ and is used as adnominal modification. The speaker explained that Muslims use a typical hat. The noun introduced by *kil* is thus not anaphoric, since *shapeu* is mentioned for the first time in the discourse. These uses of the proximal demonstrative together with *ki* in (221) are deictic.

We have examined so far the adnominal use of the distal demonstrative. Now we come to the pronominal case, which is exemplified in the sentences below:

- 222) **Kil** yera filhus de regulo i kumpanhia.  
 DEM be.PAST child-PL of king and company  
 ‘Those were the king’s sons and their fellowship.’
- 223) **Ki kwatru** mas pikininus sta na Guine.  
 DEM four more small-PL COP in Guinea  
 ‘Those younger four are in Guinea.’

In (222) *kil* has an endophoric deictic force, which refers to an antecedent in the discourse, where the speaker referred to certain persons who have a certain name. The sentence above is thus a further specification of who those people were. It is important to note that *kil* refers here to more than one individual, although it does not carry any plural specification. It could be considered as evidence of what we said at the beginning of this section, namely that Kriyol proximal and distal demonstratives do not seem to receive an overt morphological mark of plurality, even when they semantically imply plurality. The same holds for *kil* in (223), where

it is used pronominally and refers to the speaker’s siblings, whom had already been mentioned in the discourse. The proximal demonstratives in (222) and (223) are anaphoric. The longer sentence below displays first and subsequent mentions of a new discourse referent. The first mention of *dinheru* ‘money’ is introduced by *un*, which in this case seems to be a numeral rather than a determiner. The speaker is in fact expressing the quantity of coins, namely one. Subsequent mentions of *dinheru* are accompanied by *kil* in its anaphoric use:

- 224) Pa (n) yabri boka, u ten ke da **un dinheru**.  
 COMP 1sg open mouth 2sg have to give one money  
**Ki dinheru** ke e ta pudu na **kil tigela**, kada kin ke na  
 DEM money REL 3pl HAB put-ed in DEM bowl each who REL CONT  
 entra e pui dinheru. **Ki dinheru** fika pa **ki alquin** ke labau.  
 enter 3pl put money DEM money remain for DEM someone REL wash-2sg  
 ‘For me to open my mouth, you have to pay one coin. That money that was put  
 in the bowl, everyone who comes inside puts a coin (in the bowl), that money  
 is for the one who has washed you.’

Finally, as for the third use of demonstratives mentioned above, i.e. determinate reference, we already examined this case in 4.2.7.

#### 4.2.11 Agreement in Kriyol: The modifier

The lack of agreement, or at least poor agreement, seems to be something common to creoles and pidgins in general. According to Maurer et al. (2013), gender agreement is absent in most creoles and pidgins all over the world. There are, however, a number of creoles that show gender agreement between a (small) number of adjectives and the noun they modify. Among these languages, we find GBC, Cape Verdean Creole, Papiamentu, and Principense. For example, we may find in GBC AdjPs like *badjuda bonita* lit. ‘girl pretty’. This does not seem to happen among the older generation (Intumbo, Inverno & Holm 2013). When plurality is not overtly marked on the noun, the distinction between the singular and plural interpretations seems to depend, at first sight, on contextual and pragmatic cues.

In this section we will discuss a topic strictly correlated to the question of number, namely agreement. This may be defined as a “syntactic relation between words and phrases which are compatible [...] by virtue of inflections carried by at least one of them” (Matthews 2007:13). It is important to recall here that inflection is almost completely absent in Kriyol,

having as its only realization the plural marker. What one may expect from a language like that is a complete absence of agreement. Yet, the situation in Kriyol is not always straightforward, and we do sometimes find (pieces of) agreement. Before we continue, we should clarify that it does not seem to occur in the variety taken into account by Kihm (2007). It is more likely to happen in the non-basilectal grammar, i.e. the one more similar to EP. In the following description we will see two different types of agreement realization: determiners, quantifiers and modifiers may agree in number with the noun, or they may not. We already discussed the case of modifiers in 4.2.6 above. The most creole pattern is when only the noun is marked as plural, yet both noun and modifier can be pluralized. As a final case, it may happen that only the adjective is overtly pluralized. What is interesting to notice at this point is that we have two patterns of word order. Kriyol adjectives are usually postnominal; however, a small number of modifiers such as *bon* ‘good’, *diferenti* ‘different’, etc. show up in the prenominal position. In (225) the adjective *diferentes* precedes the noun it modifies, whereas *usadus* in (226) occupies a postnominal position. This is perfectly in line with what happens in EP (and more generally in Romance languages), where modifiers are normally in the postnominal position. Nonetheless, in some cases the prenominal position is preferred, e.g. the Italian adjectives *buono* ‘good’ and *bello* ‘beautiful’, which normally occur in the prenominal position.

- 225) No tene **diferentis** **etnias**.  
 1pl have different-PL ethnic.group-PL  
 ‘We have different ethnic groups.’
- 226) Li e ta bindi **karus** **usadus**.  
 LOC 3pl HAB sell car-PL used-PL  
 ‘They sell second-hand cars in that place.’

#### 4.2.12 The case of the Guinean community of Lisbon

Before we continue with the present description of the GBC nominal system, there is an interesting fact that we may want to look at regarding number agreement: among the community of Kriyol native speakers living in Lisbon, the plural marker seems to occur both on the noun and on its modifier. It seems, thus, that within this variety a kind of contact-induced form of nominal agreement is increasing, which is probably due to the daily contact that Kriyol has with its lexifier European Portuguese. This is, however, not a regular

phenomenon, and alternates with the situation described above as for agreement. In (227) and (228) we have some examples of these facts:

- 227) **Kostumes diferentes, linguas diferentes** tambe.  
custom-pl different-pl, language-pl different-pl also  
'Different customs, different languages as well.'
- 228) **Antigos combatentes livres** de luta e ta skribi ba.  
ancient-pl fighter-pl free-pl of fight 3pl HAB write ANT  
'The free ancient fighters of war, they used to write.'

These facts warrant further analysis; however, this is outside the purpose of the present work.

#### 4.2.13 The case of numerals

In this section we will deal with the case of the numeral phrase in Kriyol. What we said above about agreement between modifiers and nouns is true also in the present case: roughly speaking, Kriyol nouns do not take the mark of plurality if they are modified by a numeral or a modifier. This is also the case of quantifiers, as we mentioned above, which we will see in more detail later. The following sentences will show the behaviour of nouns in numeral phrases, i.e. in the presence of a numeral higher than one, thus implying plurality. However, we will see later that this is not the rule, but rather a tendency.

- 229) N tene **dus santchu**.  
1sg have two monkey  
'I have two monkeys.'
- 230) N ta djagasi es **dus palavra**.  
1sg HAB mix.up this two word  
'I mix up these two words.'
- 231) I **tris anu** mas bedju de ke mi.  
3sg three year more old than that me  
'He is three years older than me.'

The sentences in (229-231) all contain a numeral phrase, where the numeral implies plurality. Nonetheless, the noun introduced by the numeral does not show any plural morphology. In particular, in (229) the numeral *dus* is followed by the noun *santchu* 'monkey' in its nonplural form, although 'monkey' is an "individualizable" animal in our gradient of animacy.



Sentences (230) and (231) do not contain any human or animate noun. *Palavra* ‘word’ in (230) is introduced by the numeral *dus* ‘two’, and *anu* ‘year’ in (231) is preceded by the numeral *tris* ‘three’. Both nouns appear in their singular forms. We can conclude, thus, that number words specified for ‘more than one’ do not select for further plural specifications. Moreover, this is independent from the animacy hierarchy.

As we already said above, this state of things is merely a tendency, although it seems to be predominant. There are indeed also cases where the noun presents overt plural morphology regardless of the presence of the numeral, as the following example shows:

- 232) Na Buba un kusa de **des kasas** ke i ten ba.  
 in Buba INDEF thing of ten house-PL REL 3sg have ANT  
 ‘In Buba there were about ten inhabitations, but now there are more than one hundred.’

In the sentence (232) we see a numeral phrase that has a plural noun as its complement, i.e. *des kasas* ‘ten houses’. Although the situation sketched here does not seem to be as usual as the one described above, it seems nonetheless to confirm what we have assumed so far, namely that it is just a tendency in Kriyol, and not the rule.

#### 4.2.14 The quantifiers

The case of quantifiers in GBC will reinforce the idea that Kriyol native speakers are provided with several co-existing subgrammars. Similar to the case of the number words described above, quantifiers such as *manga di* ‘a lot of’ and *tudu* ‘all’ may select for both plural and nonplural forms. We will first examine the case of the quantifier *manga di*:

- 233) Es guera i kusa ke **manga de djinti** ta medi bai nvesti la.  
 DEM war 3sg thing REL plenty of people HAB be.afraid go invest LOC  
 ‘This war is why a lot of people are afraid of investing there.’

- 234) Pabia **manga de kuzas** ku n skribi mas n ka sta suguru  
 because plenty of thing-PL REL 1sg write but 1sg NEG stay sure  
 si i sta dritu.  
 if 3sg stay right  
 ‘Because as for many things I wrote, I am not sure whether they are right.’

In the two sentences above we have the case of the quantifier *manga di*: in (233) it takes the singular noun *djinti* as its complement, whereas in (234) the noun introduced by the quantifier is in its plural form, i.e. *kusas*. We already looked at sentence (233) for the description of demonstratives (216). We cannot address here the semantic question of animacy, introduced above, because the present case is completely reversed, since the noun [–animate] *kusa* ‘thing’ is pluralized, whereas *djinti* ‘people’, which is [+animate], does not take any plural specification. It seems that the speaker may quite freely choose between the two options.

We take now into account sentences with the quantifier *tudu*, in order to demonstrate that the situation just sketched is found not only with *manga di*, but also with other quantifiers. *Tudu* is a floating quantifier which can occur before the noun, after it, and also separated from it:

- 235) Na ki **djintis tudu**, anos tudu i guineus.  
 in DEM people-PL all 1pl.TOP all 3sg Guinean-PL  
 ‘Among all these people, we all are Guineans.’
- 236) N tene projetus manga del, mas kel **tudu projetus**, tudu bai.  
 1sg have project-PL plenty of-3sg but DEM all project-PL all go  
 ‘I had a lot of projects, but they all failed.’
- 237) **Tudu rasa** sta la.  
 all race be LOC  
 ‘All races are present there.’

In sentences (235) and (236) we find that the quantifier *tudu* does not agree with the noun it refers to. Indeed, in (235) *djintis* carries the specification of plurality, whereas *tudu* remains in its singular form, like in (236), where the plural noun *projetus* follows the quantifier. Worth noting here is that in sentence (235) the quantifier is postnominal, whereas in (236) it is prenominal. Both possibilities are found in Kriyol. In (235) we should also take note of the quantifier phrase *anos tudu*, where the first plural personal pronoun is accompanied by *tudu* in its singular form. On the other hand, the example in (237) shows a different case, i.e. the quantifier is accompanied by a nonpluralized noun, i.e. *tudu rasa* ‘all race’. The fact that the singular form of the quantifier under consideration is preferred most of the time seems to be confirmed in the example below, where the speaker started the sentence with the plural form of the quantifier *tudu*, but immediately corrected himself and decided to use the singular form:

- 238) Tudus... **tudu** pa ermon, tudu ta djuda nutru.  
 all-PL all for sibling all HAB help each.other  
 ‘All...all for one, everybody helps each other.’

### 4.3 Bare noun phrases: Distribution and interpretation

The main concern of the present section is to describe the behaviour of bare nouns in Guinea-Bissau Creole. At the beginning of the present chapter, we defined real bare nouns in GBC as NPs which have no determiner and no specification as for number. The lack of number specification will turn out to be crucial for our final analysis in Chapter 6. In other words, nominals in Kriyol are most times ‘bare’, appearing in their root form as well as Kriyol verbs. Bare nouns in this language may occur in both argumental and predicative positions. As for argument BNs, they are allowed in both the subject and object positions, i.e. external and internal objects. As for internal objects, we should distinguish between two different kinds of objects, i.e. the patient and the recipient, which in languages with ‘case’ would correspond to accusative and dative, respectively.

In the following sections, we are going to look at the distribution and interpretation of bare noun phrases. We will take into account both count and mass bare nouns, in the argument position (subject, recipient, patient), nonargument position (prepositional BNs, BNs in topic constructions, left- and right-dislocated BNs, clefted BNs, and directional objects), and BNs in the predicative position. Count BNs and mass BNs in GBC are identical in their behaviour, with the only exception that mass nouns cannot be pluralized, whereas count Ns may receive the mark of plurality. It will turn out that BNs may occur in both the subject and object position without any distributional asymmetry. As for their interpretation, BNs show a certain asymmetry. Subject BNs are more likely to yield a definite reading, and the same is true for recipient object BNs, whereas in the case of patient object BNs, no reading seems to be banned. The specific plural interpretation seems, however, to be very unlikely. Moreover, bare PPs seem to have access to any kind of interpretation. As for other types of nonargumental BNs, such as topicalized, left- and right-dislocated, and clefted BNs, they usually receive a definite interpretation. It is worth mentioning, however, that clefted BNs may also yield an indefinite interpretation. Finally, directional objects may yield any possible interpretation. As we will see in more detail in the following sections, the (in)definite and

(non)specific interpretations of GBC BNs principally depend on two factors: predicate type and the aspect of the verbal predicate (perfective or imperfective). Otherwise, the correct interpretation of a BN depends upon the context.

#### 4.3.1 Subject bare nouns

Bare subjects show a clear tendency towards the definite reading. It is important to say, however, that the indefinite reading is not excluded. We are going to examine the conditions under which a bare subject in Kriyol may yield an (in)definite and (non)specific interpretation in 4.4. For the time being, we limit our purpose to the description of subject BNs. In (239) and (240) we have examples of subject BNs in GBC: in (239) *prublema* ‘problem’ yields a definite interpretation, whereas *vida* ‘life’ in the second clause is interpreted as definite generic. Unlike in (239), the subject BN *omi* ‘man’ in (240) receives an indefinite specific reading.

- 239) **Prublema** i ke dinheru ke djintis ka ta tene na mon,  
 problem 3sg DEM money REL people-PL NEG HAB have in hand  
 I kela ke prublema ma **vida** fasil na Guine.  
 3sg DEM+LOC REL problem but life easy in Guinea  
 ‘The problem is that money that people do not have (in their hands), that is the problem, otherwise life is easy in Guinea.’
- 240) **Omi** kumpra masan aos.  
 man buy apple today  
 ‘A man bought apples today.’

The examples above have shown that both definite and indefinite interpretations are possible; however, the most usual is the definite one, as the sentences below will show.

- 241) **Renda** na karu kada byas mas.  
 rent CONT expensive each time more  
 ‘The rent is becoming more and more expensive.’
- 242) **Djenti** sta la i tera so pa lavoru.  
 people stay LOC and land only for cultivation  
 ‘People are there and the land is only for the cultivation.’

- 243) **Sol** sai.  
 sun rise  
 ‘The sun has risen.’

In example (241) *renda* ‘rent’ is the bare subject of the sentence. It receives a definite reading. In this case the interpretation is not specific, but generic; indeed, the speaker is not referring to a certain rent (of a certain house), since the discourse is about rent in general. The sentence in (242) is made up of two coordinated main clauses, each having its own (bare) subject, i.e. *djenti* ‘people’ and *tera* ‘land’, respectively. Similar to the sentence in (241), the bare subjects here receive a definite reading. In (243) the bare subject *sol* ‘sun’ receives a definite interpretation just as well as *renda*, *djenti* and *tera* in the former sentences. However, we have to notice that the definite interpretation of *sol* seems to depend on a different factor: its definiteness can be derived from the fact that the sun is a ‘unique entity’, and this seems to play a great role in terms of definiteness in other languages as well (Kihm 1994; Baptista 2007:74). In sum, the sentences in (241-243) share the definite interpretation of their bare subjects, which is not unexpected if we consider that the left periphery of the sentence is generally associated with already known information, i.e. definiteness (Chafe 1976, Krifka 2007).

So far we have examined bare subjects, which yield a definite interpretation. As we already mentioned, the definite interpretation is the most suitable for bare subjects in GBC. Nonetheless, the indefinite reading is not banned as was shown in sentence (240), which we repeat here below in (244):

- 244) **Omi** kumpra masan aos.  
 man buy apple today  
 ‘A man bought apples today.’

The bare subject *omi* ‘man’ yields an existential indefinite interpretation. The derivation of the existential reading principally depends upon the fact that *omi* is the subject of the stage-level predicate *kumpra* ‘buy’. In fact, as we will see below in 4.4, the subjects of stage-level predicates yield existential readings. Furthermore, the indefiniteness of *omi* in (244) depends upon pragmatic reasons. It is, in fact, a new discourse referent, i.e. it represents new information in the discourse context. Nonetheless, there is no doubt that bare subjects in Kriyol have a strong preference for the definite interpretation, and cases of indefinite interpretation, such as (244), seem to be limited to pragmatic conditions.

We have seen so far the behaviour of count bare nouns, and the case of bare mass nouns will turn out to be very similar. Bare mass nouns may also occur in the subject position, where they usually yield definite readings, as we will see in examples (245) and (246):

245) **Arus** tamben baratu.

Rice also cheap

‘Rice is also cheap.’

246) **Pon** bida seku.

Bread become dry

‘The bread has become dry.’<sup>80</sup>

The bare mass noun *arus* ‘rice’ in (245) yields a definite generic reading, whereas *pon* ‘bread’ in (246) is interpreted as definite nongeneric. The difference between *arus* in (245) and *pon* in (246) as to genericity arises from the different predicate types: *baratu* ‘cheap’ in (245) is an individual-level predicate which yields a generic reading; however, *seku* ‘dry’ in (246) is a stage-level predicate. That is why the bare subject *pon* receives a nongeneric (existential) reading. We will give a more detailed description of these facts in 4.4 below. However, it is interesting to note that the predicate in (245), i.e. *baratu*, is not accompanied by any copula since it is a verbal adjective. We will discuss these facts later.

As we saw above, a bare noun in a preverbal position may yield both generic and existential readings. The same is true for the kind-referring reading: GBC bare subjects may yield a kind-referring reading whenever they are accompanied by a predicate suitable for such an interpretation such as *be extinct*, *be widespread*, etc. The direct accessibility of the kind-referring reading for GBC BNs is a consequence of the fact that the basic denotation of nouns crosslinguistically is kind of type *e* (Depréz 2007). This was mentioned in 3.9 and will be analyzed in more detail in Chapter 6. For the time being, we just need to say that the kind-reading is directly accessible for BNs which are  $\neg$ [number], and also for [+number] NPs, as in example (200), repeated here below in (247). Both  $\neg$ [number] and [+number] NPs may yield the kind-referring reading; this interpretation is yielded via the kind-yielding predicate ‘be extinct’. It is important to note that the bare subject *dinossauro* ‘dinosaur’ may occur in its plural form. This is an option left to the speaker. We see here once again that requirements for overt plurality such as discourse-relevance and animacy may be overruled, in this particular case, by reference to the kind.

---

<sup>80</sup> It does not seem to be certain here that *pon* is a mass noun with the meaning of ‘bread’. Kihm (p.c.) advises that *pon* could have the meaning of ‘loaf’. This would resemble the ambiguity of EP *pāo* ‘bread/loaf’ (Kihm p.c.).

- 247) **Dinossauro(s)** ka ten dja.  
 dinosaur(-PL) NEG have already  
 ‘Dinosaurs are extinct.’

We come now to the case of the generic reading, which is different from the kind-referring, insofar it does not refer to the kind, but to (entities of) the subset of that kind, as we already discussed in Chapter 3. The bare subject *bulanha* in (248), which is an area for the cultivation of rice, yields a generic interpretation. The sentence is a description of those areas, meaning that the speaker is not referring to any particular *bulanha*. The bare subject *bulanha* is a  $\neg$ [number] bare noun, i.e. it is unspecified as to number. This is also consonant with our analysis of overt plurality by means of animacy: *bulanha* is in fact [-animated].

- 248) **Bulanha** i pertu de... pertu de riu.  
 Bulanha 3sg close of... close of river  
 ‘*Bulanhas* are close to the rivers.’

Now we come to the case of existential BNs. In (249) we have a  $\neg$ [number] BN in the subject position, which yields a plural existential reading. Following the general tendency of Kriyol as to animacy, we would expect to find an overt plural noun phrase. Once again, the animacy condition may be overruled. This is so probably because the “number” of elephants is not relevant to the discourse.

- 249) **Elefanti** ka ten la.  
 Elefant NEG have LOC  
 ‘Elephants are not to be found there.’

If we consider that Kriyol native speakers may resort to different subgrammars (Kihm 1994, 2007), the situation may turn out to be even more intricate. It seems, however, that the general tendency for the subject position is to host  $\neg$ [number] bare nouns whenever the BN yields a kind-referring or generic reading. Existential BNs may also be  $\neg$ [number] if the overt expression of plurality is not relevant to the discourse contexts. Otherwise, whenever plurality is relevant to the discourse and/or whenever the subject is an existential plural, then we generally have an overt plural noun as in (180) repeated below in (250). Finally, as for kind-referring and generic BNs, overt plurals are also possible, although the most usual case is the form  $\neg$ [number].

- 250) **Garandis** de tabanka ta pega kil, e ta tokal.  
 Old-PL of village HAB take DEM 3pl HAB play-3sg  
 ‘The old men of the village take it and play it.’

#### 4.3.2 Object bare nouns

We mentioned above that, as for bare patient objects in Kriyol, no reading seems to be banned: (in)definite, (non)specific, and (non)plural. It is nonetheless important to note that the plural specific interpretation is quite unlikely, and it is usually obtained via overt plural morphology, i.e. *-s*. The bare object *panu* ‘towel’ in (251) yields a singular indefinite nonspecific reading (251a). In (251b) we have also the definite reading of *panu*, which is not the case here, since *panu* is first mentioned and represents new information in the discourse context. Yet, in principle, the definite reading would be possible. On the other hand, the bare object *kabesa* ‘head’ in (252) is a case of inalienable possession.

- 251) U ta maradu **panu**.  
 2sg HAB tie-ed towel  
 a. ‘A towel is tied around you.’ (new referent, indefinite)  
 b. ‘The towel is tied around you.’
- 252) I ta rapadu **kabesa**, i ma raka un lesu na kabesa.  
 3sg HAB shave-ed head 3sg more tie INDEF lace in head  
 ‘His head is shaved, he also ties a lace around his head.’

Bare count nouns in the object position usually yield a reading which is unspecified for number (Kihm 2007, Depréz 2007). The bare object *panu* ‘towel’ in (253) is in fact unspecified as for number, i.e.  $\neg$ [number]. On the other hand, *kabas* ‘bowl’ in (254) is semantically interpreted as singular. Moreover, it yields an indefinite nonspecific reading.

- 253) N ta kumpra **panu**.  
 1sg HAB buy towel  
 ‘I buy towels.’
- 254) No na fitcha **imbashada**.  
 1pl CONT close embassy  
 ‘We are going to close the embassy.’



- 255) Su u na pega **kabas**, bu ten ke pegal asim.  
 If 2sg CONT pick.up bowl, 2sg have to pick.up-3sg so  
 ‘If you pick up the bowl, you have to pick it up this way.’

In principle, *panu* in (253) may yield both a singular reading and a reading unspecified for number; however, in this particular case, the right interpretation is the second one, i.e. ¬[number]. In more detail, the sentence in (253) is a generic one since it contains the habitual aspect marker *ta*. As we will see later in this chapter, the habitual aspect of the verbal predicate makes a sentence generic. In such sentences, bare objects are usually interpreted as indefinite nonspecific. The sentence was uttered during the description of a typical funeral ceremony in Guinea-Bissau, and the speaker was intending to buy items of the kind *panu*. The reference to *panu* is thus not really at stake here. In summary, sentence (253) exemplifies how the habitual aspect marker *ta* influences the interpretation of bare objects: nonspecificity is preferred, although specific readings are not banned.

Now we come to the case of the continuous aspect marker *na*, which we are going to see in more detail later in the present chapter. On the other hand, the bare object *imbashada* ‘embassy’ in (254) yields a singular specific interpretation. Specificity is yielded here on the basis of the presence of the continuous aspect marker *na*, which usually makes a sentence episodic. Episodic sentences are suitable for specific interpretations. We will see these facts in more detail in 4.4.. Finally, in sentence (255), the bare object *kabas* receives a definite nonspecific reading. Unlike example (254), the sentence in (255) is not really episodic despite the presence of *na*. The sentence, in fact, rather refers to an eventuality introduced by the complementizer *su* ‘if’. As a consequence, *kabas* in (255) yields a nonspecific reading.

Now we come to the case of bare mass objects. We already noted that bare mass Ns behave like bare count Ns. The following examples will show this:

- 256) E ta tisi bos **yagu**.  
 3pl HAB bring 2pl water  
 ‘They bring you (some)water.’
- 257) Depus ki kabra ke e ta mata, parsin kuma e ta kusinha nan **bianda**.  
 After DEM goat REL 3pl HAB kill, seem-1sg COMP 3pl HAB cook just food  
 ‘After that goat that they have killed, it seems to me that they cook more food.’
- 258) Kin ke na ponha **bianda** purmeru.  
 who REL CONT get food first  
 ‘The one who catches the food first.’

In (256), *yagu* ‘water’ is the new information, thus yielding an indefinite reading. Moreover, *yagu* is nonspecific, given the habitual context. The same holds for *bianda* ‘food’ in (257). It yields an indefinite interpretation since it represents new information in the discourse context. On the other hand, *bianda* in (258) yields a definite interpretation because the referent of the noun is already known in the context since the speaker already mentioned it. The sentence in (258) belongs, in fact, to a description of a typical wedding ceremony in Guinea-Bissau, where the bride and groom have to compete (*jumna-jumna* lit. ‘arrive.first-arrive.first’) to take the bowl of food first. Furthermore, the presence of the continuous marker *na* provides evidence for such reading.

We come now to the case of recipient objects in GBC, which usually yield a definite reading. *Mininu* ‘child’ in (259) is a bare recipient object. It directly follows the verb and precedes the direct object, i.e. patient object, which in this case is introduced by the indefinite D, i.e. *un libru* ‘a book’. The bare recipient object *mininu* yields a definite reading, whereas the patient *un libru* is indefinite. As a difference, the patient object *storia* ‘story’ in (260) is bare and receives a –[number] interpretation. The bare recipient object *mininu* is in its usual postverbal position and receives a definite reading, as in the former case.

259) Djon na da **mininu** un libru.  
 John CONT give child INDEF book  
 ‘John gives/is giving/will give the child a book.’

260) Bu ta konta **mininu** storia.  
 2sg HAB tell child story  
 ‘You tell stories to the child.’

We see, thus, that bare recipient objects are typically interpreted as definites. However, for recipient objects to yield an indefinite interpretation, they should be introduced by an indefinite D. This is shown in (261), where the recipient *mininu* is introduced by *un*.

261) Bu ta konta **un** mininu storia.  
 2sg HAB tell INDEF child story  
 ‘You tell stories to a child.’

Interestingly, the tendency for recipient objects to yield a definite reading is not exclusive of Kriyol. On the contrary, it seems to be a widespread tendency for natural languages in general. Czypionka (2014) claims that verbs that assign noncanonical nominative-dative case-patterns to their arguments “cause higher processing costs during comprehension”. In other words, recipient objects have “long-lasting effects on sentence-comprehension (in eye-

tracking and ERP measurements)” (2014:97). Moreover, Czypionka’s work concerns the fact that object-animacy also causes a longer processing in sentence comprehension. However, this is outside the purpose of the present work and should be measured with experiments such as eye-tracking. As for our present study, we should only say that (in)animacy of the recipient does not influence its interpretation as for (in)definiteness, and (in)definiteness depends upon the presence or absence of the indefinite determiner *un*.

### 4.3.3 Non-argument bare nouns

As for BNs in non-argument positions, we will look at the cases of PPs, topicalization, left- and right-dislocation, and clefting. As we will see, the behavior for possible interpretations of non-argument BNs is quite homogeneous. The definite reading is the most plausible for all of them except for PPs which may yield all possible interpretations. Moreover, clefted BNs also allow indefinite readings, although the definite one is more likely to be yielded. As a final case study for BN positions, we will take into account the directional object, which in GBC often occurs without preposition: no reading seems to be banned.

We begin with the case of PPs. Bare PPs are allowed to yield any possible interpretation; in other words, no reading seems to be banned. In (262) the PP *na prediu* (lit. ‘in building’) yields an indefinite interpretation (262a). We have to consider that, in principle, the definite interpretation is possible (262b). In this case, however, it is not the correct reading. Another issue to consider is whether the PP is specific or nonspecific, and, in principle, both readings are possible. In the present case, however, *na prediu* is indefinite nonspecific since the sentence is a generic one, and no specific building is referred to by the speaker. The same is true for the PP *ku faka* (lit. ‘with knife’) in (263). In principle, both (in)definite and (non)specific readings are possible. Importantly, it seems that both definite and indefinite readings of *faka* in (263) are more likely to be specific, since the sentence is episodic and perfective.

- 262) Kin ku mora por esemplu **na prediu**, i ka fasil.  
 Who REL live for example in building 3sg NEG do-3sg
- a. ‘Who lives for example in a building, he does not do it.’  
 b. ‘Who lives for example in the building, he does not do it.’

- 263) Maria lanha mon **ku faka**.  
 Mary cut hand with knife  
 ‘Mary cut her hand with a/the knife.’

As for topicalization, BNs in such a position usually yield a definite reading. The topicalized BN *kumboju* ‘train’ in (264) is in fact a definite generic NP.

- 264) **Kumboju**, no ka tene kumboju.  
 Train 1pl NEG have train  
 ‘As for trains, we do not have trains.’

The case of left-dislocated BNs is not different from topicalized BNs. They yield a definite reading. The left-dislocated BN *fruta* ‘fruit’ in (265) receives a definite generic interpretation.

- 265) **Fruta** tambi i ten na Guine manga de.  
 fruit also 3sg have in Guinea plenty of  
 ‘(As for) fruit, there is plenty of it in Guinea.’

The same is true for right-dislocated BNs. As we said previously about topicalization and left-dislocation, right-dislocated BNs also yield a definite reading. In (266) we have a right-dislocated BN *situason* ‘situation’, which is modified by an AdjP. It is interpreted as definite.

- 266) Situason ka sta la dritu, **situason** propi ekonomika di dinheru.  
 situation NEG stay LOC right situation proper economic of money  
 ‘The situation there is not good, (I mean) the economical situation.’

Finally, we have the case of clefted structures, which, similar to the cases analyzed so far, also yield a definite reading. Yet, this is not the only possible interpretation. The indefinite one is not banned. In (267) the clefted BN *mindjer* ‘woman’ is a case of unique referent in the discourse context where the sentence has been uttered. In this case, the noun yields a definite reading.

- 267) I son **mindjer** ke na sustenta kasa.  
 3sg only woman REL CONT support house  
 ‘It is only the woman that supports the house.’

Before we conclude this section, we have to mention the case of directional objects, which are usually expressed via BNs in Kriyol. In (268) *fera* ‘markt’ yields a definite interpretation. The indefinite interpretation is not banned altogether.

- 268) N dibi di bai **fera** aos.  
 1sg have to go market today  
 ‘I have to go to the market today.’

The sentence in (268) contains a directional object without preposition. In this same position we could expect a PP in languages such as English, German and Italian<sup>81</sup>. However, the Kriyol verb of direction *bai* ‘go’ does not seem to require any preposition. Kriyol seems to use the preposition *pa* with place names. The example below will clarify this fact:

- 269) Djalo i bai dja **pa Guiné**.  
 Djalo 3sg go already to Guinea  
 ‘Djalo has already gone to Guinea.’

The case in (269), where the place name is introduced by the directional preposition *pa* seems to be evidence of the usual behavior of Kriyol proper names in the position of directional objects (Kihm 1994). However, the presence of a bare noun as directional object is not a rule in Kriyol. There are in fact also cases where the preposition *pa* accompanies the noun, as the sentences below clearly shows:

- 270) N dibi di bai **pa fisioterapia**.  
 1sg have to go to physiotherapy  
 ‘I have to go to the physiotherapy office.’
- 271) N bai dja nan ku nha familias **pa matu**.  
 1sg go already just with my parent-pl to wood  
 ‘I had already gone into the forest with my family.’

#### 4.3.4 Bare nouns as predicates

We already mentioned that BNs in Kriyol may occur in both argument and non-argument positions. Among the non-argument positions in which a BN may appear, we also find the predicative one. GBC BNs in the predicate position are very common and may yield both singular and plural interpretations. In the present section, we will observe all the cases mentioned above: predicate BNs with singular readings, predicate BNs with plural readings and predicate plural NPs.

---

<sup>81</sup> There are some exceptions in English, e.g. *go home* vs. *\*go to home*.

As for predicate BNs with singular readings, we have some examples in (272) and (273). The bare predicates *peskadur* ‘fisherman’ in (272) and *poeta* ‘poet’ and *politiku* ‘politician’ in (273) yield a singular indefinite interpretation. They “agree” with their singular subjects *kil omi* ‘that man’ in (272) and the third singular pronoun *i* ‘s/he/it’, respectively.

272) Kil omi i **peskadur**.

DEM man 3sg fisherman

‘That man is a fisherman.’

273) Paralel di musiku i sedu tambe **poeta**, depus i **politiku**.

parallel of musician 3sg COP also poet after 3sg politician

‘Apart from being a musician, he was also a poet, and finally a politician.’

One distinction is that the sentence in (274) has a plural reference, i.e. *kil omis* ‘those men’. Nonetheless, the predicate BN is in its usual singular form *peskadur*, like in (272). It is important to note that the predicate may also occur in its overt plural form *peskaduris* ‘fishermen’. In other words, overt pluralization of predicate BNs is optional. As for the subject, it seems that overt pluralization is required. For example, the sentence in (272) has only one reading, namely the singular one. I assume that the same requirements for overt plurality that we saw in 4.2.6 are at play here, i.e. discourse relevance, animacy and referentiality.

274) Kil omis i **peskadur(is)**.

DEM man-PL 3sg fisherman(-PL)

‘Those men are fishermen.’

There is an interesting difference between (272) and (274), on the one hand, and (273), on the other. Sentences (272) and (274) do not seem to contain any overt copula. Instead, we find the third singular pronoun *i* as ‘link’ between the subject and noun predicate. This is in contrast to (273), where the copula *sedu* ‘be’ occurs. Interestingly, the pronoun *i* also plays the function of copula in Kriyol. It happens with noun predicates and certain adjective predicates (real adjectives) in present- and past-tensed copular sentences. We will see these facts in more detail below. The copula *sedu* seems to occur only in certain contexts, namely “with less basic, stylistically more or less ‘learned’ quality items” like *demokratiku* ‘democratic’ (Kihm 1994:35; emphasis in original). According to Kihm (2007b), we should not define *i* as a copula since it shows nonverbal behavior. In more detail, whenever the negation *ka* co-occurs with *i* in a copular sentence, *ka* follows *i*. Differently, *ka* always precedes the verbs as we saw as to negative declarative sentences in Chapter 2. On the basis of this non-complete verbal behavior of *i* in copular sentences, Kihm assumes *i* to be a predicate marker, and not a full copula.

Now we come to the case of noun predicates introduced by the indefinite determiner. In (275) the noun predicate is the indefinite NP *un mediku* ‘a doctor’, and not a bare noun as in the cases discussed above. The presence of *un* seems to depend on the presence of the relative clause that modifies the noun predicate. The indefinite determiner probably has the function of yielding the correct reading of the modified noun predicate, i.e. the indefinite reading. If *mediku* had not been introduced by *un*, the most felicitous reading for it would be the definite one (276). As we saw in 4.2.7, the demonstrative *kil* does not necessarily appear in such contexts. The bare noun *mediku* in (276) is perfectly felicitous.

275) El i yera ba **un mediku** ku djintis gosta ba  
 3sg.TOP CIS-3sg be.PAST ANT INDEF doctor REL people-pl like ANT  
 del tchiu.  
 of-3sg much  
 ‘He was a doctor who people really liked.’

276) El i yera ba **mediku** ku djintis gosta ba del tchiu.  
 3sg.TOP 3sg be.PAST ANT doctor REL people-PL like ANT of-3sg much  
 ‘He was a doctor who people really liked.’

We have seen so far that BNs in the predicate position yield an indefinite reading, if not modified. At this point an interesting question may arise: may non-modified predicative BNs also yield a definite interpretation? Although the definite interpretation is not banned altogether, the indefinite one is much more felicitous. When it comes to expressing definiteness, Kriyol speakers preferentially use constructs like (277) which is the answer to hypothetical questions such as *Kin ki mediku?* ‘Who is the doctor?’ and *El i ki?* ‘Who is s/he?’.

277) El ki mediku.  
 3sg.TOP REL doctor  
 ‘S/He is the doctor.’

We have so far discussed GBC BNs in argument, non-argument and predicate positions. We saw that BNs in this creole have a very free distribution and may yield all types of possible readings, yet with certain differences in relation to the syntactic position where they occur and the thematic role they get. Before we discuss the factors involved in the derivation of the correct interpretation of BNs in Kriyol, we need to briefly consider predicative sentences in GBC and its copula(s). Indeed, predicative sentences and the copula system are one further aspect where GBC consistently differs from its lexifier EP.

#### 4.3.5 The copula system of Guinea-Bissau Creole: An overview

Kriyol has a rich copular system. It varies depending on the type of predicate and tense of the sentence. Whenever the copular sentence contains a noun predicate, the third singular pronoun *i* will connect the subject and the predicative noun like in (272) and (274). In some cases, the overt verbal copula *sedu* occurs instead of *i*. This principally happens with less basic items such as *poeta* ‘poet’ in (273) and adjectives such as *demokratiku* ‘democratic’, as we mentioned above. We already mentioned that Kihm (2007b) assumes *i* to be a predicate marker and not a full copula. As for predicative adjectives, we observe the same alternation as for noun predicates: *i* and *sedu* compete for the same function, although the typical case is that where *i* occurs, as we already discussed. There is also an interesting alternation between overt and null copula as to adjectival predicates, but this is outside the aim of the present work. We only need to observe that there is in Kriyol a class of quality items (adjectives) that show verbal behavior (Kihm 1994). This entails that no copula introduces them. These verbal items are property items referring to “basic qualities like big, small, happy, sad” (1994:34), which syntactically behave like verbs. This means that aspect markers *na* and *ta* may precede them. Also, the tense markers *ba* (postverbal) and *na bin* (preverbal) may accompany a verbal adjective. For instance, we may find sentences such as *I na garandi*, 3sg CONT big, ‘S/He is growing up’. Moreover, in case a pronoun is the syntactic subject, these verbal items obligatorily require the clitic form as their subject (see discussion above in Chapter 2 on GBC clitic pronominal system). For example, a sentence like ‘I am happy’ would translate into Kriyol *N kontenti* ‘1sg happy’ and not into *\*Ami kontenti* ‘1sg.TOPIC happy’.<sup>82</sup> For a more accurate description of predicative adjectives in Kriyol, see Kihm (1994, 2007b).

Before we turn to the issue of tense in copular sentences, we need to briefly discuss the item *i*. The choice of *i* is not surprising, if we consider that *i* in Kriyol is not only the third singular clitic pronoun, but also a kind of “neutral pronoun analogous to English *it* or, better still, French *ça*, which is able to refer back to anything in the world” acting as a kind of “universal antecedent” (Kihm 1994:160f.).

---

<sup>82</sup> We already discussed that the clitic subject is the real syntactic subject. Its presence is mandatory in languages that necessarily require an overt subject in order to satisfy the EPP requirement. In GBC, an instance of such clitic subject is 2sg *bu*, often phonetically reduced to *u*, like in *bu kontenti* ‘you are happy’. If *kontenti* were a real adjective, *bu* would not be required but rather *abo*, the strong (optional) pronoun which is typically a focusing element, would be required, like in many other cases of clitic doubling.



As for tense, whenever a copular sentence is in the present tense, the pronoun *i* in its copular function occurs. In some cases, we find *sedu* instead of *i* (see above). This is exemplified in (272) and (274), on the one hand, and in (273) on the other. This is true for both nominal and adjectival predicates. However, if the adjectival predicate belongs to the class of verbal items, no copula shows up (see below). Other contexts where *sedu* shows up are in the presence of an aspect marker and following complementizers. As for the former case, *sedu* appears when the character of the sentence is expressed via an aspect marker. In other words, when we have a generic copular sentence with the habitual marker *ta* or an episodic copular sentence with the continuity marker *na*, *sedu* will follow the aspect marker (278).

- 278) Ka ta sedu suma bu madrinha.  
 NEG HAB COP like your godmother  
 ‘(The mother of the wedding) is not like your godmother.’

As for embedded copular sentences introduced by complementizers, the presence of *sedu* is usually required (279).

- 279) Dipus e ta tisi bos utru kabas ke sedu ki de kbamba.  
 Then 3pl HAB bring 2pl.OBL other bowl REL COP DEM of kbamba  
 ‘After that, they bring you another bowl, which is the one with *kbamba*’.

Copular sentences in the past tense show the same alternation observed as for copular sentences in the present tense. The copular item (*i* or *sedu*; in the case of verbal adjectives, no copula is required) must be accompanied by the past marker *ba*. As a difference, however, Kriyol also displays a verbal copula specialized for the past tense, i.e. (*y*)*era*. This past copula may be accompanied by *ba*, as in (275) above. In (280) we have a past-tense copular sentence containing a verbal adjective. As a consequence, no copula will show up, and the past is expressed by the marker *ba* only. Finally, in (281) we have the past copula *yera* without further specification of past via *ba*.

- 280) Kasa ku kai bedju ba di mas.  
 house REL fall old ANT very  
 ‘The house that fell down was very old.’
- 281) N yera pikininu, inda n ka garandi.  
 1sg be.PAST little still 1sg NEG big  
 ‘I was still a young boy, I was not an adult yet.’

At the level of microvariation, we find another interesting case in nonpresent tense predicative sentences in GBC. Indeed, the copulas *é* and *foi*, present and past (perfective) Portuguese copulas respectively, may be occasionally found.

- 282) E la ke foi nha obietivu.  
DEM LOC REL COP.PAST.PERF my objective  
'This (very one) was my objective.'

Kriyol did not take them as copulas, at least not during the stage of creolization. I would not say that a new copula system is emerging. Rather, it seems that it is a parallel system available to the speaker, but it is still understood as *mas portuguis* ('more Portuguese') by the speakers. Interestingly, there is no copula specified for posteriority, i.e. for future tenses, in this parallel system; at least, I did not find it yet.

Finally, future-tense copula sentences always contain a copula. In fact, *sedu* always shows up and is preceded by a future-specifying marker (or construct): *na sedu* or *na bin sedu*. The former one may indicate future or progressivity exactly as the continuity marker *na* does. On the other hand, the latter one indicates the specific future.

- 283) Bu fidju na sedu matchu.  
Your child CONT COP male  
'Your child will be a boy.'
- 284) Alguin ka sibi inda ke ku na bin sedu dipus.  
some NEG know still what+REL spec.fut. COP after  
'We do not know what will happen.'

For the sake of completion, we must note that Kriyol has another copula, namely *sta* 'be, stay'. It derives from EP *estar* 'be, stay' and is the locative copula. *Sta* also has the function of expressing temporary states predicated of (sets of) individuals. In other words, *sta* is used in combination with adjectives or adverbs in order to form stage-level predicates, e.g. *sta duenti* 'be sick', *sta in dia* 'be up-to-date, be in good standing', etc. Finally, *i* sometimes exhibits the function of a locative copula.

## 4.4 Derivation of BN interpretation

In the present section, we will look at the factors which contribute to the derivation of the correct interpretation of BNs in Kriyol. Both syntactic, semantic and pragmatic conditions should be taken into account. Beside the context, which is sometimes crucial as to the correct interpretation of BNs, factors such as verbal aspect (perfective vs. imperfective) and predicate type (individual-level vs. stage-level) play a crucial role.

### 4.4.1 How aspect influences BN interpretation

Progressive and perfect relate to events and typically trigger episodic (particular) readings of sentences. Sometimes they also show up in generic (characterizing) sentences (Krifka 1995). This is a description that can be applied in principle to any language; it does not regard Kriyol only. In (285a) and (285b) we have a progressive and a perfect sentence, respectively. Both sentences are episodic.

- 285) a. John is smoking a pipe.  
b. John has smoked a pipe.

Habitual aspect, temporal adverbs like *usually* and auxiliary constructions like *used to* typically yield generic interpretations (ibid.). In (286a) we have a simple present, in (286b) a simple past, and in (286c) a frequency adverb. All three sentences are habitual.

- 286) a. John smokes a pipe.  
b. John smoked a pipe.  
c. John is often smoking a pipe.

As for BNs in the object position, they yield a definite reading in perfective contexts. On the other hand, if they occur in imperfective contexts, they are interpreted as indefinite (Krifka 1989). We will see this in more detail in Kriyol, but we first need to look at aspect and tense in this creole.

As we already said, verbal aspect is relevant to the interpretation of BNs in Kriyol. GBC bare forms of dynamic verbs yield a simple perfective reading, but bare forms of stative verbs represent “presently obtaining states of affairs” (Kihm 2011:92). As we explained in Chapter 2, Kriyol has two imperfective preverbal aspect markers: the habitual *ta* and the continuous *na*. The continuous *na* may be either progressive or not, and may also be habitual

(Comrie 1976). Adverbs such as *normalmenti* ‘usually, generally’ or *sempri* ‘always’ may also contribute to the generic character of a sentence.

As we mentioned above, bare objects in past perfective contexts yield definite (specific) readings. Yet, the indefinite reading is not excluded; the context must be suitable. The bare object *tesi* ‘thesis’ in (287) receives a definite specific interpretation. *Skola* ‘school’ in (288) is also a definite specific BN, since it is definite in context.

- 287) N kaba **tesi** na janeiro sim de es anu, n terga **tesi**.  
 1sg finish thesis in January yes of DEM year 1sg submit thesis  
 ‘I finished the thesis last January, I submitted the thesis, now I am waiting to defend it.’
- 288) E bin bai rebenta **skola**.  
 3pl POST go make.explode school  
 a. ‘Then they made a school explode.’  
 b. ‘Then they made the school explode.’

Similarly, bare objects are also interpreted as definite (specific) in imperfective (continuous) contexts with *na*. Both bare objects *imbashada* ‘embassy’ and *bilheti di passage* ‘tickets (of journey)’ in (289) yield a definite specific reading. The same is true as for the bare object *luta* ‘fight’ in (290).

- 289) Bom, no **na** fitcha **imbashada** na Berlin, no **na** mandabos **bilheti di**  
 well 1pl CONT close embassy in Berlin 1pl CONT send-2pl ticket of  
**passage** ke dinheru pa fitcha imbashada na Berlim.  
 journey with money to close embassy in Berlin  
 ‘Well, we are closing the embassy in Berlin, we will be sending you the tickets together with the money for closing the embassy in Berlin.’
- 290) I sta konvinsidu kuma i **na** ganha **luta** tambai pa bai luta  
 3sg stay convince-ed that 3sg CONT obtain fight also to go fight  
 son pabia di konfiansa.  
 only because of trust  
 ‘He is sure that he is also going to win the fight, (thus) he is going to fight just because of self-confidence.’

It is interesting to note that the continuous aspect marker *na* does not always yield episodic readings of the sentence and, more importantly, specific readings of the BN. The sentence in (255), repeated below in (291), represents a special case of nonepisodic-inducing *na* due to

the presence of the complementizer *su* ‘if’. As a consequence, the bare object *kabas* ‘bowl’ is interpreted as nonspecific.

- 291) Su u na pega **kabas**, bu ten ke pegal asim.  
 If 2sg CONT pick.up bowl, 2sg have to pick.up-3sg so  
 ‘If you pick up the bowl, you have to pick it up this way.’

As a last case of aspect triggering a certain BN interpretation, we will now consider the habitual context with the aspect marker *ta*. In such contexts, bare objects yield an indefinite nonspecific interpretation. Both BNs in the object position *tarbadju* ‘work’ in (292) and *pis* ‘fish’ in (293) are indefinite nonspecific.

- 292) Bu **ta** otcha me **tarbadju** ma utrora bu ka ta risibi dritu.  
 2sg HAB find INTERJ work but sometimes 2sg NEG HAB receive right  
 ‘You find a job, ok, but sometimes you do not receive good pay.’
- 293) Bu **ta** kume **pis** la ku un euro.  
 2sg HAB eat fish LOC with one euro  
 ‘You can eat fish there with just one euro.’

#### 4.4.2 Predicate types and the interpretation of bare subjects

Predicate types are relevant to the interpretation of bare subjects, as was noted by Carlson (1977) regarding English bare plural subjects. Carlson distinguishes between stage-level (294a) and individual-level predicate types (294b):

- 294) a. (The) Firemen are available.  
 (stage-level predicate, temporal property)
- b. Firemen are altruistic.  
 (individual-level predicate, permanent property)  
 (Kratzer 1995: 125)

The subjects of stage-level predicates yield existential readings, either definite or indefinite. On the other hand, subjects of individual-level predicates usually receive a generic interpretation. The same can be said for GBC subject BNs. We will first examine the case of bare subjects in the presence of stage-level predicates. The bare noun *elefanti* ‘elephant’ in (295) is the subject of a stage-level predicate, i.e. *(ka) ten* ‘are (not) (there)’. It yields an

existential indefinite reading. On the other hand, the bare subject *prublema* ‘problem’ in (296) is in the presence of the stage-level predicate *tene na mon* ‘have in hand’ and is interpreted as existential definite.

295) **Elefanti ka ten la.**

Elephant NEG have LOC

‘There are no elephants there.’

296) **Prublema i ke dinheru ke djintis ka ta tene na mon.**

problem 3sg DEM money REL people-PL NEG HAB have in hand

‘The problem is that money that people do not have (in their hands).’

As a difference, GBC bare subjects in the presence of an individual-level predicate yield a definite generic reading. The BN *bida* ‘life’ in (297) is the subject of the individual-level predicate *i baratu* ‘is cheap’, and yields a definite generic reading, as explained above. The same is true for the BN *konvivensia* ‘living together’ in (298). It is the subject of the individual-level predicate *i diferenti* ‘is different’ and, as expected, it gets a definite generic interpretation.

297) **Bida la i muito baratu.**

life LOC 3sg very cheap

‘Life there is very cheap.’

298) **Konvivensia, ke manhera ku no ta sta la djuntu sim**

living.together DEM manner REL 1pl HAB stay LOC together so

ku nutru, i diferenti ke li.

with one+other 3sg different REL LOC

‘Living together, the way we live all together with each other, it is different from here.’

## 4.5 Summary and conclusion

In the present chapter, we supplied an exhaustive description of the Kriyol nominal system. We took into account both its NPs and functional items. We have furthermore considered the crucial issue of overt plurality in GBC. We saw that relevance of plurality in the discourse is the first condition for overt pluralization of Ns. Furthermore, animacy and referentiality also represent crucial conditions.

Importantly, we mentioned that BNs in Kriyol are unspecified for semantic number, which will turn out to be crucial for our theoretical implications, the topic of Chapter 6. After we gave a brief definition of BNs, we described GBC BNs in argument, non-argument and predicate positions. It turned out that there is no subject/object asymmetry for BN distribution. As for their interpretation, BNs show a certain asymmetry. The subject demonstrates a strong preference for the definite reading. The same can be said for the recipient object, which always yields a definite reading. On the other hand, object BNs may yield any possible reading except for the plural specific. It is in fact quite unlikely that a BN yields such a reading. The overt plural would fit better.

Moreover, we said that Kriyol has no definite determiner, which implies that definiteness is expressed via BNs. As for indefiniteness, we have in GBC an indefinite determiner, i.e. *un* ‘a/an’. This does not always show up in indefinite contexts. This means that indefiniteness is also realized via BNs. As for *un*, it usually expresses nonspecific indefiniteness.

Finally, we considered the factors that influence the correct interpretation of GBC BNs. They are aspect and tense as for bare objects, on the one hand, and predicate types as for bare subjects, on the other. In more detail, we saw that habitual aspect triggers indefinite readings for object BNs, and perfective and continuous aspects trigger definite readings. On the other hand, as for bare subjects, we found that individual-level predicates trigger generic interpretations, whereas stage-level predicates trigger existential (in)definite readings. Among other issues, we must make a more detailed analysis of specificity issues as for indefinite BNs and *un*-NPs, and prosody and phrasing with respect to BNs.

## Chapter 5

### Crosslinguistic comparison of bare nouns

#### 5.1 Cape Verdean Creole

Cape Verdean Creole (hence CVC) is one of the Upper Guinea Creoles, in addition to Guinea-Bissau Creole and Ziguinchor Creole<sup>83</sup>, and shares with them Portuguese as a lexifier language. It is spoken on the nine islands of Cape Verde, which are located off the coast of Senegal. The Cape Verdean islands are traditionally divided into two groups, i.e. the one in the south, called Sotavento (or Leeward Islands), and the other in the north, named Barlavento (or Windward Islands). CVC splits into more varieties, i.e. that of Santiago, Brava, and São Vicente (Lang 2013, Baptista 2013, and Swolkien 2013). However, we will look in more detail at the varieties described in Baptista (2002, 2007), namely the ones spoken on the Sotavento Islands of Brava, Fogo, Santiago and Maio. Before continuing with the description of CVC bare nouns, it is worth noting the high similarity between CVC and GBC. In fact, they are closely related to each other. It is possible that the very same Portuguese Pidgin acting as the basis for the creole of Guinea-Bissau also served as the basis for CVC (Kihm 1994).

##### 5.1.1 The nominal and determiner system of CVC

Like in GBC, there is no gender distinctive marker in CVC, at least not in its determiner system. According to Baptista (2013), although nouns are usually invariable, nominal expressions of the type *matxu* ‘male’ and *femia* ‘female’ may appear to the immediate right of nouns denoting “animals, kinship terms, names of professions, or properties characterizing humans” (Baptista 2013:14). In other words, we find in CVC the very same animacy hierarchy acting in GBC.

According to Baptista (2002, 2007), CVC has two overt determiners: the indefinite *un* and its plural counterpart *uns*. They come from the EP masculine indefinite determiner *um*

---

<sup>83</sup> Although Ziguinchor Creole seems to be simply a variety of GBC, some linguists prefer to describe it as an independent creole.



and its plural counterpart *uns*. This closely resembles the case of GBC, where the very same indefinite determiner is present. In CVC, *un* may also be a numeral and a quantifier. Baptista (2002) assumes *un* to be (non)specific. Furthermore, *uns* is better described as a quantifier than as a real determiner. The examples below will show the above described facts: in (299) *un* is an indefinite determiner and yields a nonspecific reading. In (300) we have an instance of specific reading of *un*, and in (301), *un* occurs in its quantifier role. Finally, in (302) the plural *uns* is a quantifier as well and corresponds to the English ‘some’.

- 299) Si’ n atxa meiu di n sai pa **un** kau, ma ben inviste nan ha tera.  
 If I find way of I leave for a place COMP come invest in my country  
 ‘If I find a way of leaving for a (different) place, I will return to invest in my country.’<sup>84</sup>
- 300) Nos e un povu ki ka ta xatia-s.  
 We COP INDEF people that NEG TMA bother-them  
 ‘We are people who don’t bother-them.’ (from Baptista 2002:25)
- 301) E bota-m **un** ajudinha.  
 She throw-me Q little help  
 ‘She gave me some little help.’
- 302) Es faze **uns** batankinha di midju.  
 They made some rolls of corn  
 ‘They made some corn rolls.’

As for the demonstrative system, CVC has two demonstratives, the singular *kel* and its plural counterpart *kes*. More specifically, these are the forms for adnominal demonstratives in CVC. As remarked in Baptista (2013:14), pronominal demonstratives are *es li* ‘this one’, *keli/kila* ‘this one/that one’, and *kes li/kes la* ‘these ones/those’. They derive from the EP distal demonstrative *aquela*, *-s* ‘that, those’. Sometimes the adnominal demonstrative in CVC is used as a definite determiner. This also happens in GBC, as we already saw. The facts described so far are not surprising. According to Lyons (1999) and Baptista (2007), it is quite common that definite determiners derive from demonstratives and indefinite determiners from quantifiers (singular numeral). The path of derivation would be one of semantic weakening. In CVC, the demonstratives *kel* and *kes* may mark definiteness when they refer to a previously mentioned entity (anaphoric use); however, sometimes they precede a noun even if no mention has been made formerly in the discourse. In these cases, they may yield a specific or a nonspecific reading, depending on contextual cues. The sentences in (303) and (304)

---

<sup>84</sup> If not otherwise specified, the examples in this section are taken from Baptista (2007).

provide evidence of the specific and nonspecific interpretations of demonstratives, respectively. *Kel* in (303) is used anaphorically, but not in (304).

- 303) E ten guarda kel lutu. (Lutu has been formerly mentioned in the discourse)  
 He has keep the mourning  
 ‘He must keep the mourning.’
- 304) N ta labanta nha, tres ora madrugada, n po kel midju na pilon.  
 I TMA get up woman three hour early morning I put DEF corn in pestle  
 ‘I get up, woman, at three o’clock in the morning. I put the corn in the pestle.’

With respect to plural marking, CVC has adopted different strategies. An element may be placed before the noun in order to yield a plural reading. More specifically, Baptista assumes that the elements that have a pluralizing function are the plural indefinite determiner (quantifier) *uns*, the plural demonstrative *kes*, possessives - specified for plurality via the plural suffix *-s* (e.g. *nha-s*, ‘my-pl’), numerals, or floating quantifiers such as *tudu* ‘all’. Since each of the above mentioned pluralizing elements are specified for plurality, either inherently or via the plural suffix, the noun is not necessarily marked as plural and thus may remain in its (unmarked) bare form. In (305) we find the plural demonstrative *kes*, which precedes the noun *fidju* ‘son’ in its unmarked form. In (306), the numerals *seti* ‘seven’, *sinku* ‘five, and *dos* ‘two’ introduce nouns which are not specified for number either.

- 305) **Kes** fidju, es san a ses kaza kuk ses fidja.  
 These children they are in their house with their children  
 ‘These children, they are in their houses with their children.’
- 306) N ten **seti** fidju, **sinku** matxu, **dos** femia.  
 I have seven child five boy two girl  
 ‘I have seven children, five boys two girls.’

These pluralizing strategies are partially different from those of GBC, where possessives cannot receive the mark of plural. Moreover, the demonstratives *kil* and *es* do not receive any plural marker, thus the plural is either specified on the noun or derived via a bare noun, as we already saw in the preceding chapter. This last case, i.e. plurality via bare nouns, is also present in CVC:

- 307) **Kaza** d’es rua ta parse bedju.  
 House of+this street TMA look old  
 ‘The houses in this street look old.’

Finally, CVC also displays an overt way of marking plurality on the noun, namely the plural suffix *-s*. This closely resembles the situation found in GBC, where plurality is expressed by the plural marker *-(i)s*. As we saw in the case of plurality in GBC, animacy plays a crucial role in determining whether a noun is going to be overtly marked as plural or not. The very same holds for CVC, as the following example will show (307). According to Baptista, furthermore, a noun marked for plurality may also be introduced by the plural demonstrative *kes*. However, this seems to happen rarely.

- 308) Nha **fidjus** tudu ta papia ku el.  
 My children all TMA talk with him  
 ‘All my children talk to him.’

Interestingly, plurality in CVC can be expressed on a single element of the NP, more specifically on the determiner (or even demonstrative or possessive), letting the noun unmarked for plurality (309). However, if the noun is [+human], it will be preferentially overtly marked as plural together with the determiner (310), as shown in the following example from Ferreira (2009:113f).

- 309) kes partidu ki sta faze pulitika  
 DET party that TMA do politics  
 ‘These political parties’ (my translation of Ferreira’s original in Portuguese)
- 310) nhas irmons  
 my-pl sibling-pl  
 ‘my siblings’

Further options for marking plurality call for a numeral or a quantifier such as *txeu* ‘many’. Before we conclude this section on the nominal system of CVC, we will briefly describe the behavior of adjectives in this creole. Like in GBC, adjectives are postnominal in CVC. Only a small group of adjectives inherited from EP seems to be prenominal. The modifier *bon* ‘good’ in (311) occurs in prenominal position, whereas *spertu* ‘intelligent’ in (312) occurs after the noun.

- 311) Nu fase un **bon** trabadju, pa nu pode ten midju.  
 We did a good work for we can have corn  
 ‘We did good work so that we could get corn.’

- 312) Un mininu **spertu** ta trabadja riju.  
A child intelligent ASP work hard  
'A bright child works hard.'

### 5.1.2 Bare nouns in CVC

Although CVC does not have a poor determiner system, as we saw above, bare nouns in this language seem to occur quite freely. Their occurrence is, however, limited by some constraints, or rather tendencies. As we will see in the present section, we find quite strong tendencies acting in CVC and regulating the occurrence of bare nouns, which will result in a situation not very different from what we saw in GBC. Let us see that in greater detail.

According to Baptista (2007), in CVC bare nouns and NPs introduced by the demonstrative *kel* in its role of determiner seem to occur in free variation. The very same holds for the indefinite determiner *un*. These facts seem to provide evidence for the fact that in CVC, overt DPs of the type *kel* + N and *un* + N do not necessarily yield specific readings. On the contrary, these kinds of DPs may also be nonspecific. As we saw in the preceding chapter regarding the indefinite determiner *un*, the situation seems similar to the one in GBC. More specifically, *un* usually yields a nonspecific interpretation, although we may also find cases where the specific reading is derived. On the other hand, as for GBC demonstrative *kil* in its determiner function, it behaves differently from its CVC counterpart and always yields specific readings. However, in CVC generic NPs seem to be expressed via bare singular nouns only (Baptista 2013).

With respect to bare nouns, they may be found in CVC both in the subject and object positions. However, depending on their syntactic position, they yield different interpretations. Baptista (2002, 2007) assumes that BNs are subject to a high degree of interpretational variation. They may be interpreted as generic, indefinite plural, indefinite singular, definite singular, or definite plural, depending on "contextual, contextual cues or simply pragmatics" (2002:30). More in detail, BNs may yield a definite specific reading in the subject position, either singular or plural. On the other hand, when they occur as objects, they receive an indefinite reading, either specific or nonspecific. Like in GBC, these are just tendencies. Although strong, they allow a certain range of variation. The sentences in (313-317) show the behavior of bare subjects and a bare object. *Kaza* 'house' in (313) receives a definite specific plural interpretation. *Milágru* 'miracle' in (314) and *makáku* 'ape' in (315) also receive a plural

reading, but not specific. They are, indeed, generic (note that also the sentences where they occur are generic). On the other hand, the bare singular objects *terenu* ‘terrain’ in (316) and *txafaris*<sup>85</sup> ‘well’ in (317) are interpreted as singulars. More specifically, *terenu* yields an indefinite nonspecific reading, whereas *txafaris* is assigned a definite reading.

- 313) **Kaza** di es aldeia e baratu.  
House of this neighborhood is cheap  
‘The houses in this neighborhood are cheap.’
- 314) **Milágru** [...] ka ten o si e ten nos tudu nu pode fase-l.  
Miracle Neg have or if 3sg have 1pl.INDEP all 1pl.DEP can do-3sg  
‘No such things as miracles [...] exist, or, if they do exist, then we all are able to perform them.’ (Veiga 1987:69 in Baptista 2013)
- 315) **Makáku** tioxu ka tonba, purki es ê kustumádu ku rótxa.  
Ape never Neg fall because 3pl be familiar with rock  
‘Monkeys never fall down because they are used to rocks.’  
(Lang 2002 in Baptista 2013)
- 316) Si bu ten **terenu** di fase, bu ta pidi djuda.  
If you have terrain to make you TMA ask help  
‘If you have a terrain to make, you can ask for help.’
- 317) Nu ten **txafaris** pa nu panhu agu.  
We have well to we take water  
‘We have the well to collect water.’

With respect to overt plurals, Baptista further assumes that overt plural nouns in CVC yield definite specific or indefinite (non)specific interpretations both in the subject and object positions. In (318) below, the plural object *fidjus* ‘children’ receives a definite specific interpretation.

- 318) Ta bende un balai di pexi, nu ta kre pa kria **fidjus**.  
TMA sell a basket of fish we TMA want to raise children  
‘We sell a basket of fish, we need it to raise the children.’

There is a further distinction to be made, i.e. between count and mass nouns. Following Baptista, countable BNs can be interpreted as generic, (in)definite specific plural, indefinite

---

<sup>85</sup> Note that *txafaris* ‘well’ is a singular form, although it ends in *-is*. However, in this case *-is* is not the plural suffix. In fact, *txafaris* comes from EP *chafariz*, *-es* ‘well, -s’.

nonspecific plural, indefinite (non)specific singular, or definite (non)specific singular, as stated above. On the other hand, mass nouns such as *agu* ‘water’ are typically bare. Hence, they do not show the same degree of variation found between bare count nouns and their determined counterparts. Finally, bare mass nouns freely occur both in the subject and object positions. The example below in (319) is a repetition of (317) above. This time I highlighted in bold the mass noun *agu*.

- 319) Nu ten txafaris pa nu panhu **agu**.  
We have well to we take water  
‘We have the well to collect water.’

In summary, singular BNs are usually interpreted as definite or indefinite depending upon their position in the sentence. More specifically, they yield a definite reading when they occur in the subject position, whereas bare singular objects are usually assigned an indefinite reading. The subject position is, in fact, related to older information, whereas what is new is usually expressed in the object position. Also worth noting is that it seems that singular bare nouns do not yield specific readings. This is only possible with bare plural nouns, which may get a specific or a nonspecific interpretation according to the context (Baptista 2002:32). We should recall that these are just tendencies: although strong, they are not the rule.

In conclusion, “there are several ways why bare NPs may appear. First, the speaker may consider information regarding specificity as irrelevant. Second, in the realm of definite NPs, the entity may be easily identifiable by both listener and speaker if such an entity is familiar to their world” (Baptista 2002:31). Last, when a NP has already been introduced by the indefinite article, its second occurrence may be bare.

## 5.2 The Creole of São Tomé

Together with Angolar, Lung’ie and Fa d’Ambô, Santome is one of the Portuguese-based creoles of the Gulf of Guinea. It is spoken on the island of S. Tomé in the Democratic Republic of São Tomé and Príncipe, located in the Gulf of Guinea, West Africa. The official (and most widely spoken) language is Portuguese; nonetheless, “Santomé is spoken and understood by most of the population, which includes Principense and Angolar speakers” (Hagemeijer 2013:50). It seems that its use among younger speakers is declining since the

number of Santome speakers has not grown. As already mentioned above, Santome is a Portuguese-based creole and, among its substrate languages, it counts several Benue-Congo languages such as Edo and Kikongo.

### 5.2.1 The nominal system of Santome

In Santome there is no morphological gender distinction, “but there are cases where the morphological contrast found in Portuguese has lexicalized”, i.e. *manu* ‘brother’ vs. *mana* ‘syster’ (Hagemeijer 2013:52). Like in the case of GBC and CVC, gender distinction for animates (in this case, nonhumans) rely on a lexical choice: *ome* ‘male, man’ and *mwala* ‘woman, female’, are inserted at the immediate right of the noun to distinguish between a male and female animal, e.g. *bwê ome* ‘bull’ vs. *bwê mwala* ‘cow’.

Unlike the Upper Guinea creoles analyzed above, Santome has a definite determiner able to act as the definite article ‘the’ in English. More precisely, Santome has a definite determiner specified for plurality, i.e. the pronominal determiner *inen*, but without a singular counterpart. According to Alexandre & Hagemeijer (2007), *inen* is, thus, inherently specified for both definiteness and plurality. As the authors remark, the very same form *inen* is syncretic with the third plural personal pronoun, which could explain why plurality is inherently specified. Therefore, *inen* can be translated either as ‘they’ or as ‘the’ in its plural reading. As for its origin, Alexandre & Hagemeijer state that it is still unclear: *inen* may derive either from Edo *iran* ‘they’ (both weak and strong personal pronoun), Edo (*n*)*ene* ‘the’ (definite article), or Edo *ene/enena* ‘these’ (demonstrative). It may also be used in order to address “a number of hearers with the meaning of ‘you’ ” (2007:41). Furthermore, in its determiner function, *inen* is pronominal, thus introducing the noun it ‘determines’. Consonant with the fact that Santome does not have any kind of nominal inflection for number, the noun introduced by *inen* will not be overtly marked as plural. It seems, thus, that Santome requires overt pluralization on the determiner only. These facts led the authors to assume that “number is parasitic on definiteness” (2007: 37). Moreover, if no other modifier occurs, the plural definite determiner *inen* introduces nouns that are [+human], a fact that is explained by Alexandre & Hagemeijer as “a consequence of its sensitivity to a semantic principle that takes [+ human] as being more individuated than items that are [-human]” (2007:41). As we will see later, if another modifier is present, *inen* may also introduce nouns that refer to inanimates. *Inen* may, furthermore, allow for gapping. It is worth noting that this definite

determiner seems to preferentially show up together with other modifiers such as possessives, relative clauses or with other “items contributing to specificity”, as remarked by Alexandre and Hagemeyer. The following sentences provide evidence for what was said so far about *inen*, showing, furthermore, that nouns introduced by the definite determiner may occur both in the subject and object positions:

- 320) **Inen funxônariu** ska bi golo pixi blatu ô.  
 DEF.PL functionary ASP go search fish cheap EMPH  
 ‘The employees are searching for cheap fish.’
- 321) Ola ku mina be, mina da ku **inen mosu**.  
 When KU girl go girl give with DEF.PL boy  
 ‘When the girl left, she ran into the boys.’

In (320), *inen* introduces an NP in the subject position, whereas in (321) it modifies a noun in the object position, and there is no difference in the resulting interpretation. This fact may be interpreted as further evidence that definiteness and number are inherently specified in *inen*. Thus, it is not surprising that this determiner is not sensitive with respect to its external or internal object position. As a crucial difference, it should be noted that, when *inen* co-occurs with other modifiers such as numerals, it yields a different reading depending on its position with respect to the other modifier, but this seems to depend on the position of the numeral, and not of the determiner.<sup>86</sup>

With respect to indefiniteness, Santome patterns along with GBC and CVC in having an overt indefinite determiner, i.e. *ũa* ‘a, an’, also derived from the Portuguese indefinite determiner *um, -a* ‘a/an’. Unlike Upper Guinea creoles, Santome derived its indefinite article from the feminine form, and not from the masculine. *ũa* is also a quantifier, and as quantifier it has a plural counterpart, although this is not its “direct” counterpart: *ũa dôsu* ‘some’ (lit. ‘one two, a couple’). As remarked by the authors, the counterpart of a NP introduced by *ũa* is a bare noun, as we will see in more detail later.

---

<sup>86</sup> As noted in Alexandre & Hagemeyer (2007:45), “numerals may occur in several different syntactic positions”, and depending “on where they occur with respect to the NP and to the D, they yield different interpretations. I will supply here the same examples as the authors, in order to show these facts:

- a. *Inen dosu mosu se*.  
 DEF.PL two boy SP  
 ‘The two boys.’
- b. *Dosu inen mosu se*.  
 ‘Two of the boys.’
- c. *Inen mosu se dosu*.  
 ‘These/those two boys.’



- 322) **Ūa soya** glavi ku mwala glavi.  
 A story nice with woman pretty  
 ‘A nice story with pretty women.’
- 323) N tē **ūa mosu** ku **ūa mina** mwala.  
 1sg have a boy and a child woman  
 ‘I have a boy and a girl.’

The sentence in (322) contains a noun introduced by the indefinite determiner *ūa* in the subject position. In the sentence in (323), on the other hand, two nouns modified by *ūa* show up in two lexically governed positions, i.e. internal object and prepositional object respectively.

Another important issue to be discussed with respect to bare nouns and, more in general, nominal systems, is the semantic category of specificity. As we already saw above, languages usually express specificity, which can be done in several ways. For instance, Mauritian Creole, a French-based creole spoken on the island of Mauritius, uses a postnominal specificity marker *là* (see Guillemin 2011). Similarly, in the case of Santome we find a postnominal specificity marker *se*. Its origin is not certain. Alexandre & Hagemeyer contemplate two possibilities: it derives either from the Portuguese demonstrative *esse* ‘that’ or from the contraction of *sa ai* ‘is here’. Continuing with our discussion on *se*, we should note that it does not allow for gapping, i.e. it always requires a nominal host. It is, thus, analyzed as clitic by Alexandre and Hagemeyer. It is worth mentioning, here, that other scholars prefer to describe *se* as a demonstrative, which may also act as definite determiner (Ferraz 1979, Lucchesi 1993, among others). Lucchesi (1993) further questions the role of *se* as a determiner, claiming that in its definite use, this item just emphasizes the noun. I will follow here the definition given in Alexandre and Hagemeyer (2007) and consider it as a specificity marker. Indeed, as remarked by the authors, *se* is unlikely to be a definite determiner on the basis of the possibility for it to co-occur both with definite and indefinite determiners. Unlike *inen*, which typically occurs with other modifiers, as we said above, *se* always shows up without any other modifier in its syntactic (postnominal) slot. Whenever the specificity marker modifies a noun, it yields a singular definite interpretation. On the other hand, if the noun marked as specific by *se* is also modified by *inen*, it will receive a plural reading. As remarked by the authors, N + *se* may also derive a taxonomic reading, if the noun is [-human]; however, a generic reading is excluded. Finally, worth noting is that the occurrence of *se* with nouns does not depend on their [ $\pm$ human,  $\pm$ animate] feature: it is, thus,

a further difference from the plural definite determiner *inen*, which occurs with [+human] nouns only.

- 324) **Mwala se**, Zon pag'e.  
 Woman SP Zon pay-3sg  
 'The woman, Zon payed her.'
- 325) N mēsê pa men jê mina **awa se da anzu se**.  
 1sg want for mother fetch little water SP give baby SP  
 'I want you to bring a bit of water (in question) for the baby (in question).'<sup>87</sup>
- 326) Inen dôsu **mosu se**...  
 DEF.PL two boy SP  
 'The two boys...'

In (324), *se* marks for specificity a noun with the feature [+human], whereas in (325) we find two occurrences of *se*: the first one with a [-animate] noun, and the second one with a [+animate] noun. Furthermore, the sentences in (324) and (325) show that the nouns marked for specificity may occur both as external argument and in lexically governed positions. Finally, (326) shows the possibility of co-occurrence of more markers in the same DP: more precisely, the definite plural determiner, a numeral and the specificity marker co-occur to modify the noun *mosu* 'boy' which, thus, receives a plural definite specific reading.

Before concluding this section on the nominal and determiner system of Santome, we should briefly discuss the Santome demonstrative system. In this creole, there are different forms of demonstratives, depending on whether the entity referred to by the demonstrative is close or out of sight. According to Hagemeyer (2013:52), there are the following forms: *ise* 'this, these', *ixi* 'that, those' (out of sight), *isaki* 'this, these' and *isala* 'that, those'. Finally, it is worth noting that "distance contrasts are often expressed periphrastically" in Santome (ibid.). If the entity in question is out of sight, the demonstrative *xi* is inserted after the noun. These facts are illustrated in the following examples (from Hagemeyer 2013:52):

- 327) Ke (se ku) sa ai/ala.  
 House DEM REL be here/there  
 'This/that house'.

---

<sup>87</sup> As remarked by Alexandre & Hagemeyer, *men* in (273) is used in an affective way to address a woman (2007:44).

- 328) Kume xi ku ê kume.  
 food DEM REL 3sg eat  
 ‘The food s/he eats’.

### 5.2.2 Bare nouns in Santome

Like the other creoles described so far, Santome makes quite free use of bare nouns, i.e. noun phrases not marked for number and not introduced by any determiner. In this creole, bare nouns may occur in all syntactic positions. Furthermore, they may show up as both arguments and predicates. It is important to note that, as in the cases of Guinea-Bissau and Cape Verdean creoles, Santome bare nouns cannot receive the whole range of possible interpretations. Let us look at these facts in more detail.

Bare noun phrases in terms of null determiner and unmarked for number are very widespread in Santome because of the absence of an overt way of marking plurality on the noun. Plurality is, indeed, marked by *inen* (see above), which, however, also marks the noun with respect to definiteness. The indefinite plural interpretation seems thus to be left out from the system of (overt) determiners in Santome. It is not surprising that this is the interpretation typically given to bare nouns in lexically governed positions. However, it is important to recall here that the same interpretation is yielded by *ũa dôsu* ‘some’. The sentence in (329) below, where the bare nouns in the object position yield a plural indefinite reading, clearly shows the tendency for the bare object to be interpreted as plural indefinite. In this respect, Santome does not differ from the other natural languages with respect to information structure. The object position is, in fact, typically associated with new information, and likewise, new information is associated with indefiniteness.

- 329) Zon bê **mwala/bisu/floli** ni losa.  
 Zon see woman/bird/flower at plantation  
 ‘Zon saw women/birds/flowers at the plantation.’

The bare nouns in (329) provide evidence for what was said above: when a bare noun carries new information, it is typically interpreted as indefinite and plural. Alexandre and Hagemeyer (2007) give two further translations of the sentence in (329). In the first one, *mwala*, *bisu*, and *floli* are interpreted as singular definites, whereas, in the second one, these bare nouns receive a plural definite reading. In both cases, the bare nouns have to be previously licensed. In other words, lexically governed positions do not always give rise to indefinite plural readings.

Indeed, if the bare noun in the object position had been previously licensed, it will yield a definite reading, either singular or plural. The number interpretation will also be previously licensed together with definiteness. The sentence in (330) also contains a bare noun introduced by a preposition, i.e. *ni losa* ‘at (the) plantation’. It receives a definite singular interpretation. Moreover, there are two exceptional cases, which turn out to be always singular, i.e. *sangê* and *sungê*, formal for ‘lady’ and ‘mister’, respectively.

- 330) Zon bê **sangê/sungê** ni losa.  
 Zon see lady/mister at plantation  
 ‘Zon saw the lady/the man at the plantation.’

On the other hand, bare subjects are tendentially interpreted as definites: the definite reading seems to be generally available, but, nonetheless, a definite reading of bare nouns is more likely to be previously licensed in the discourse world. Moreover, it is worth noting that they could occur in the sentence-initial position independent of whether they refer to humans, animates or inanimates. In fact, bare nouns referring to entities occupying both high and low positions in the gradient of animacy are felicitous as topics in Santome. However, we should notice that there is a difference in the interpretation with respect to number. Bare nouns with the feature [+human] are more likely to be interpreted as singular, although both singular and plural interpretations are available. On the other hand, [–human] bare nouns are usually interpreted as plurals.

- 331) **Mwala** kopla pixi.  
 Woman buy fish  
 ‘The woman/women bought fish.’
- 332) **Floli** môlê.  
 Flower die  
 ‘The flower(s) died.’

So far we have seen where a bare noun is allowed to occur, and which interpretation it may yield depending on its syntactic position, its human or nonhuman status, and its (eventual) previous licensing in the discourse world. Now, it is worth mentioning further contexts where bare nouns are felicitous. As remarked in Alexandre & Hagemeijer (2007), typical environments for the occurrence of bare nouns are generic sentences (333) and light verb constructions (334). Also in such contexts, bare nouns in Santome show the behavior described so far, and tend to be interpreted as singular definites in subject positions, and plural

indefinites in object positions. Furthermore, in generic sentences, the bare nouns will be generic as in (333).

- 333) Maji **vinpema** ka fe **ome** mali.  
But palm-wine ASP do man bad  
'But palm wine is harmful for men/a man.'
- 334) Sun **da** pema **faka**.  
Mr. give palmtree knife  
'The man cut an opening in the palmtree.'

In sum, what we have seen for bare nouns in Santome is not very different from the cases of GBC and CVC. Also in these Upper Guinea creoles, in fact, bare nouns in the subject position tend to be interpreted as definites, whereas lexically governed positions preferentially yield indefinite readings. A first difference lies in the number interpretation. Since both GBC and CVC have plural suffixes to mark plurality on the noun (without the consequence of further marking the noun for definiteness; see above for the discussion about *inen*), the bare noun (unmarked for number) will be tendentially interpreted as singular with both humans and animates. The speaker has, in fact, an overt plural form in order to refer to plural entities. As for inanimates and some animates (without particular cultural value), they usually yield a plural reading. On the other hand, this is not true in the case of Santome. A noun cannot be overtly marked for plurality without recurring to the definite plural determiner. Therefore, subject bare nouns with the feature [-human] will be interpreted as plurals, whereas subject [+human] bare nouns are preferentially interpreted as singulars, given their more individualizable status.

### 5.3 Papiamentu

As described in Kouwenberg and Murray (1994), Papiamentu is a creole language spoken on the so-called ABC Caribbean islands of Aruba, Bonaire, and Curaçao with a total of almost 300,000 speakers (see Kramer 2005). Another 30.000 Papiamentu speakers live in the Netherlands. Apart from the case of Aruba, where Papiamentu became the official language in 2003 (Kouwenberg 2013), the official language of the islands is Dutch, not Papiamentu. Nonetheless, Dutch is used principally for education and administration, as well as Portuguese in Guinea-Bissau. Unlike Guinea-Bissau Creole, Papiamentu does not play the

role of a lingua franca on the islands. In fact, their inhabitants “consider themselves polyglots, with varying levels of competence in Papiamentu, Dutch, Spanish and English” (1994:4). Papiamentu crucially differs from the other creoles analyzed so far in having more European languages as superstrate. It is often referred to as a Spanish/Portuguese Creole, with a (smaller) part of the lexicon coming from Dutch. English may be considered as a further lexifier language “in particular in the area of technological innovation [...]; it is to be expected that the number of words of English etymology will continue to grow and expand in other areas” (Kouwenberg and Murray 1994:6). Most scholars agree that the first lexifier was Portuguese, or rather they argue for an early Portuguese creole, as remarked by Kouwenberg and Murray, followed by Spanish and, last, by Dutch. However, others such as Maduro (1966) assume that Spanish played a more important role. Kramer (2005) also belongs to the “pro-Portuguese”, which provide an analysis of Papiamentu words that clearly have Portuguese structure, although their form seems to have come closer to Spanish in a second moment. According to Holm (1992), Papiamentu “reveals unmistakable Portuguese influence in its most basic vocabulary” (41). Following Holm, the Portuguese initial influence “has been obscured by the creole’s three centuries of close contact with the Spanish spoken by the inhabitants of nearby Venezuela” (ibid.).

### 5.3.1 The nominal system of Papiamentu

As for natural gender, Maurer (2013:166) claims that Papiamentu has two ways of distinguishing male and female entities, depending on whether they are humans or nonhuman animates. In the former case, namely with [+human] nouns, Papiamentu postpone *hòmber* ‘man’ or *muhé* ‘woman’, e.g. *yu hòmbèr* ‘son’ vs. *yu muhé* ‘daughter’. As for nonhuman animates, while for female animates the gender-indicating word is still *muhé*, male inanimates behave differently and require *machu* ‘male’ as the second word of the compound, e.g. *un buriku macho* ‘a male donkey’ vs. *un buriku muhé* ‘a female donkey’.

Kouwenberg and Murray (1994) seem to classify two items as articles, i.e. the singular indefinite determiner *un*, which we found in all the creoles analyzed so far, and the singular definite article *e*. Important to note here is that, if we consider *e* as a singular definite determiner corresponding to English ‘the’, Papiamentu turns out to be crucially different from the other Portuguese creoles taken into account in the present study. More specifically, GBC, CVC and Santome do not show any form for the singular definite article. On the other hand,

for its plural counterpart, we found the item *inen* in Santome. However, although *e* seems to be very similar to Kriyol proximal demonstrative *e(s)*, we should better not hypothesize that the form *e* in Papiamentu derives from Portuguese proximal demonstrative *este* ‘this’, as in the case of Kriyol. Scholars like Kouwenberg & Ramos-Michel (2007) assume that the Papiamentu definite article *e* comes from the Spanish singular definite article *el* ‘the’ (2007:323). Moreover, *e* in Papiamentu is syncretic with the third singular pronoun and is distinct from the demonstrative. Concluding the present discussion on the category ‘definiteness’ in Papiamentu, Arends et al. (2006) also list this semantic category among the overt inflectional items in Papiamentu (2006:225, Table 2).

The second determiner individuated in Papiamentu is *un* ‘a/an’, which is syncretic with the numeral *un* ‘one’. Kouwenberg and Murray (1994) describe *un* as [-generic, -plural, -definite], whereas *e* is [-generic, +definite]. As a confirmation of this fact, Kester and Schmitt (2007) assume that the definite determiner *e* does not allow kind readings. This interpretation is, indeed, only licensed by a bare noun, as we will see below. Furthermore, *e* never occurs in generic contexts. In sum, the definite determiner in Papiamentu only appears in nongeneric contexts, independent of whether the noun is count or mass, as the ungrammaticality of *e* in (335) and (336) will show. Furthermore, a noun introduced by *e* should be anchored in the discourse. This explains why *wil* ‘wheel’ in (337) cannot be introduced by the definite determiner. Rather, the bare noun *wil* occurs, since it is not anchored in the context (2007:119f.).

- 335) (\*E) *leon ta biba na Afrika.*  
 DEF lion PRES live in Africa  
 ‘The lion lives in Africa.’
- 336) (\*E) *lechi ta sano.*  
 DEF milk PRES healthy  
 ‘Milk is healthy.’
- 337) *Niun hende no sa ken a inventa \*e wil/wil.*  
 No person not know who PAST invent the wheel/wheel  
 ‘Nobody knows who invented the wheel.’

The definite article in Papiamentu, as described in Kester and Schmitt (2007), may have a unique reference to a previously mentioned object in the discourse via identity or part-whole relations, and allow for both referential and attributive uses (2007:117). Furthermore, Kouwenberg and Murray (1994) describe the cases where the locative adverbs *aki* ‘here’, *ei*

‘there’, and *aya* ‘yonder’ follow a definite NP, as marked for proximity. We found a very similar situation in Guinea-Bissau Creole, where in cases of proximity, the NP is introduced by the demonstrative *e(s)* and followed by the locative adverb *li* ‘here’. Whenever the distal demonstrative *ki(l)* introduces the NP followed by the locative *li*, instead of the proximal *e(s)*, what is at stake is (anaphoric) specificity, and the NP must be somehow anchored in the discourse. An example of NP marked for proximity in Papiamentu is provided here below (from Kouwenberg and Murray 1994:37):

- 338) E pòtrèt aki a wordo saká dor di e mucha hòmber ku mi ta duna  
 The picture this Asp PassAux taken by of the child male that 1sg Asp give  
 lès merdia nan.  
 lesson midday Pl  
 ‘This picture was taken by the boy whom I teach middays.’

Maurer (2013:168) assumes that the demonstrative system is a three-way distance contrast, and that the demonstratives are made of *e* (definite article) and *akí* ‘here’, *ei* ‘there’ and *ayá* ‘yonder’ at the end of the NP:

- 339) E tipo ku gusta usa kuchú akí.  
 DEM guy REL like use knife DEM  
 ‘the guy who loves to handle knives’

Concluding this section on demonstratives, the demonstrative pronouns are *esaki*, *esei*, and *esaya*.

As for the indefinite determiner *un*, Kester and Schmitt (2007) assume that it may have both narrow and wide scope:

- 340) Maria kier kasa ku un brasileño.  
 M. want marry with a Brazilian  
 ‘Maria wants to marry a Brazilian.’

The sentence in (340) seems to allow for both a narrow and wide scope reading, namely for a nonspecific and a specific reading of *un brasileño* ‘a Brazilian’, whereas a bare noun in place of the NP introduced by *un* in (340) could receive only a narrow scope (i.e. nonspecific) reading only, as we will see later.



Furthermore, although *un* seems to be preferentially nongeneric, it may yield a quantificational generic reading (341). Moreover, it may receive a kind reading but only if taxonomic (342):

341) Un hulandes ta papia hulandes.

INDEF Dutch PRES speak Dutch

‘A Dutchman speaks Dutch.’

342) #Un kabritu ta komún na Kòrsou.

INDEF goat PRES common in Curaçao

‘A goat is common in Curaçao.’

It is important here to note, following Kester and Schmitt, that *un kabritu* in (342) is odd as well as its English counterpart *goat* as subject of a predicate, which typically applies to kinds, i.e. *be common*. In order to make the sentence more acceptable in both languages, we should modify the NP through an adjective or a part-whole *with* construction, e.g. *A cat with yellow eyes is rare*.

Another category to be analyzed within the nominal system of Papiamentu is plurality. Similar to the other Romance-creoles analyzed so far, Papiamentu has an overt way to mark plurality, namely the clitic *nan*. It is suffixed to the noun to be pluralized whenever it is definite. *Nan* could also be in enclisis on a modifier of the NP to be pluralized, e.g. on an adjective. Important to note, here, is that the adjective may be prenominal or postnominal, as assumed in Kouwenberg and Murray (1994:48). Thus, it seems that only certain adjectives such as *delaster* ‘last’, *promé* ‘first’, *henter* ‘whole’, etc. may occur in prenominal position. We found a similar situation in Guinea-Bissau Creole, where adjectives are normally in postnominal position, but for some exceptions such as *bon*, ‘good’, etc., which occur prenominally.

Turning to the plural marker in Papiamentu, whenever the plural noun is generic, it does not need to be modified by *nan*, at least not in the object position. But again, whenever we have a modified NP (e.g. modified by a possessive), plurality will be overtly marked. Kouwenberg and Murray describe these facts as crucially depending on definiteness (1994:49). Worth noting is that, in the data provided by Kester and Schmitt (2007), the bare nouns that are semantically plural but not overtly marked as such, are in the object position and could be analyzed as lexically incorporated in the verb phrase.

*Nan* is described as a phrasal clitic by Kouwenberg and Murray (1994:49f.) since it may be in enclisis either on the noun (head noun of a NP) or at the end of a NP (e.g. on an adjective), as noted above. They furthermore recall what Dijkhoff (1983) said about *nan*, namely that it can also modify pronouns and other pronominal forms such as the demonstrative in its pronominal use and relative pronouns (1994:50). The sentence in (343) contains the demonstrative *es* marked for plurality: *esnan* ‘those’ (from Kouwenberg and Murray 1994:39).

- 343) Ta solamente esnan ku no tin plaka.  
 Be only those that not have money  
 ‘(It’s) only those that do not have money.’

Similar to the Santome case, where the plural (definite) element *inen* is syncretic with the form of the third plural pronoun, *nan* in Papiamentu is a syncretic form for both the plural marker and the third person plural pronoun *nan* (Kouwenberg and Murray 1994).

As for the category agreement, as remarked by the authors, there is no agreement in the creole under study. Similarly, case distinction is not morphologically marked. In fact, “the order of constituents reflects the syntactic relations. It is strictly SVO, and the indirect objects precede the direct objects” (1994:35). The considerations given to the plural marker in Papiamentu made so far lead to the conclusion that, according to Kester and Schmitt (2007), bare plurals in Papiamentu are crucially restricted in their occurrences. Indeed, since *nan* is specified for definiteness, overt plurals are always definite. As said above, *nan* is indeed specified for definiteness. Thus, whenever an NP is followed by this plural marker, it will be automatically specified for definiteness, and not only for plurality.

### 5.3.2 Bare nouns in Papiamentu

As a commonality to the other creoles analyzed above, we may find both singular and (overt) plural bare nouns in Papiamentu. According to Kester & Schmitt (2007), bare singulars in this creole would behave exactly like bare plurals in English. They could, indeed, occur “as subjects of predicates that apply to kinds only” (2007:110). Furthermore, their interpretation varies between a generic and an existential reading. The derivation of the correct interpretation would depend on the type of predicate used. Lastly, as assumed by the authors, if elements such as negation, intensional verbs such as ‘want’, and durative

adverbials are present in the same sentence where a singular BN occurs, then the bare noun allows a narrow scope reading only. The sentences below in (344) and (345) show the first two properties of singular BNs in Papiamentu as described above. More precisely, the sentence in (344) contains a predicate that occurs exclusively with names of kinds, and is a generic sentence. Hence, the BN *kabritu* ‘goat’ receives a generic reading. On the other hand, in (345) the bare singular *kas* ‘house’ is interpreted as existential and varies between a singular and a plural reading, depending on the context.

- 344) Kabritu ta mashá komun na Kòrsou.  
 Goat is very common in Curaçao  
 ‘Goats are very common in Curaçao.’
- 345) Mi a kumpra kas.  
 1sg PAST buy house  
 ‘I bought a house/houses.’

As for the scopal properties of the singular BNs in Papiamentu, we already said that the narrow scope reading is the only possible if certain elements are present, i.e. negation, durative adverbials, or intensional verbs. This is shown in (346) and (347), where the bare nouns *mancha* ‘spot’ and *brasileño* ‘Brazilian’ are nonspecific.

- 346) Mi no a mira mancha riba suela.  
 1sg not PAST see spot on floor  
 ‘I didn’t see spots on the floor.’
- 347) Mi a mata yuana pa dos ora largu.  
 1sg PAST kill iguana for two hour long  
 ‘I killed iguanas for two hours.’

As remarked by Kester & Schmitt, the bare singular *yuana* ‘iguana’ in (347) yields a plural interpretation. In order to derive an indefinite specific singular reading, the indefinite determiner *un* should introduce the NP, but the result would be odd since “we have to imagine a very special scenario of perhaps being able to kill the same iguana many times” (2007:112).

According to Kester & Schmitt, bare singulars in Papiamentu may furthermore show up in contexts of inalienable possession (348) and with unique and familiar entities (348). In the latter case, we should specify, as remarked by the authors, that the bare singular in these contexts is felicitous even if there is no previous mention of the entity in question. This seems, however, not to be characteristic of Papiamentu. In fact, unique entities such as the sun

or the moon occur typically as bare nouns in many creoles, independently from a previous mention in the discourse world. As we will see below, the situation is quite different in languages such as Brazilian Portuguese, which requires an overt definite determiner in contexts of inalienable possession, with unique entities, and generally whenever there is no previous mention of the entity picked up by the NP.

Turning to the present analysis of bare nouns in Papiamentu, take a look at the following sentences. Examples (348) and (349) demonstrate inalienable possession and unique entity in Papiamentu, respectively.

348) El a laba (su) kara.

He PAST wash (his) face

‘He washed his face.’

349) Solo ta kima sin misericòrdia.

Sun PRES burn without mercy

‘The sun is burning without mercy.’

In the sentence in (348) the bare singular *kara* ‘face’ may show up with or without the possessive, which is not surprising. In many languages there is no need for possessives in constructions of inalienable possession such as body parts. For instance, the same sentence in Kriyol would have the very same two options: either a bare singular or a NP introduced by a possessive: *I laba (si) kara*, lit. ‘S/He washed his/her face’. However, in languages such as Italian, inalienable possession is usually expressed without possessives, but with the definite determiner and a verbal reflexive construction: *Mi lavo la faccia*, lit. ‘I wash myself the face’.

In (349), the bare noun *solo* denotes the unique entity ‘sun’. It does not need to be introduced by a definite article, as would be the case in languages such as English or Romance.

As for bare plurals in Papiamentu, Kester & Schmitt note that plural BNs have a much more restricted occurrence than singular BNs. As we saw above, bare singulars may occur in both the subject and object positions, and yield both generic and existential readings (under certain circumstances, i.e. depending on other elements present in the sentence). The same does not hold for bare plurals in Papiamentu, which are felicitous as subjects in episodic sentences only. Hence, bare plurals do not yield a generic reading. On the other hand, if the sentence is generic, then the bare subject must be singular. These facts are shown in the sentences below; in (350), we find an episodic sentence with the bare plural subject *hendenan*

‘persons’, which is felicitous in such a context. On the other hand, (351) represents a generic sentence, where the bare plural *hendenan* is not allowed. Instead, its singular counterpart is felicitous.

- 350) Despues ku \*hende/hendenan a keha, nan a drecha e pelicula.  
 after that person/person-pl PAST complain they PAST fix the film  
 ‘After some people complained, they fixed the film.’
- 351) Si hende/\*hendenan ta keha, no wòri ku nan.  
 If person/\*person-pl PRES complain not worry with them  
 ‘If people complain, don’t worry about them.’

Furthermore, bare plurals in Papiamentu do not seem to allow a generic reading neither as subjects nor as objects, as shown in the following examples, which contain the bare plural subject *muchanan* ‘children’ and the bare plural object *pushinan* ‘cats’, respectively.

- 352) \*Muchanan ta inteligente.  
 Child-pl PRES intelligent  
 ‘Children are intelligent.’
- 353) \*Mi ta gusta pushinan.  
 1sg PRES like cat-pl  
 ‘I love cats.’

As for the object position, bare plurals are licensed only if modified. If they are not modified, they are not allowed as internal arguments. This makes the authors conclude that bare plurals in Papiamentu yield a specific reading, in the sense that they cannot yield kind readings and must be anchored in the discourse (2007:116). The different behavior of a bare plural object and a modified NP in the object position is shown in (354) and (355), respectively. In greater detail, the sentence in (354) shows the ungrammaticality of the bare plural object *bukinan* ‘books’. Instead of it, the speaker will use its singular counterpart *buki* ‘book’. On the other hand, in (355) the bare plural *bukinan* is felicitous as the object, since it is modified by *na spañò* ‘in Spanish’.

- 354) Mi ta mira \*bukinan/buki riba mesa.  
 1sg PRES see book-pl/book on table  
 ‘I see books on the table.’

- 355) Mi ta mira bukinan na spañò riba mesa.  
1sg PRES see book-pl in Spanish on table  
'I see Spanish books on the table.

In sum, we noticed an asymmetry in the distribution and interpretation of bare singulars and bare plurals in Papiamentu. More specifically, bare singulars are allowed as both external and internal arguments of the verb. As for their interpretation, they may yield both generic and existential reading, basically depending on the predicate that accompanies them. Furthermore, the distribution of bare plurals is much more restricted if compared to that of bare singulars. Indeed, they may occur as subjects of episodic sentences only, and they are never allowed in generic contexts. Moreover, bare plural objects are allowed only if modified. This further entails that bare plurals cannot yield generic readings, neither as subjects nor as objects. These facts lead Kester & Schmitt to the conclusion that they are inherently specific.

## 5.4 Brazilian Portuguese

It is important to distinguish between European Portuguese and Brazilian Portuguese, since there are many grammatical differences between the two varieties. As for Brazilian Portuguese (henceforth BP), Holm (1992) distinguishes standard vs. nonstandard varieties. The Popular Brazilian Portuguese (PBP) is the nonstandard variety, "spoken by lower-class Brazilians with little education" (1992:37), whereas Standard Brazilian Portuguese (SBP) is "the literary language usually spoken by middle and upper class Brazilians" (ibid.). As listed in Holm's study, among the most striking divergences between the two Portuguese varieties with respect to morphology, we find the loss of number agreement within the NP and the lack of agreement between the verb and its subject. A first explanation for the lack of number agreement in the nominal domain of PBP is phonotactic in nature. Following Holm, there is a tendency for the syllable in PBP to be open of the type CV, and this prevents the plural suffix *-s* to be realized, e.g. PBP *nas selva* (SBP *nas selvas*), 'in the forests' (49). However, as already remarked by Holm, this alone is not sufficient to explain this phenomenon, and it is important to notice that the loss of number agreement is quite widespread in Atlantic creoles (49). This brings us into the heart of the typological discussion on BP, and explains why the case of BP was inserted in the 'creole' part of this chapter, and not in the section on bare nouns in noncreole languages. Indeed, many scholars consider BP as a semi-creole. To better

understand this point, let me briefly review the major approaches to BP typology and language history.

As for the origin of this language, Holm assumes BP to be a semi-creole. In Holm's words:

*“Popular Brazilian Portuguese did indeed evolve out of the contact between a noncreole (the colloquial and regional Portuguese brought from Europe to Brazil from the sixteenth century to the present) and a creole (the Portuguese-based variety brought from São Tomé to Brazil during the sixteenth and seventeenth centuries, and later varieties such as Helvécia Creole Portuguese”*<sup>88</sup> (62).

Baxter also assumes a “creolistic” position. In his study of Helvécia creole (1992), principally based on data regarding subject-verb agreement, he claims that “em Helvécia houve uma criouliização” (‘in Helvécia, there was a creolization’ - my translation) – (1992:30). Guy (1989) and Lucchesi (1998), among others, also defend the creole hypothesis.

On the other hand, there are positions that radically differ from the “creole” one. For instance, Ferreira (2009) assumes a ‘variationist perspective’ based on sociolinguistic variation, proposing different scenarios for the linguistic variation of BP, especially as for the variation of nominal and verbal agreement. However, she remarks that BP has an African heritage, like many creole languages, most probably from the sub-Saharan region. She agrees with Holm (1992) in listing the loss of number agreement among the grammatical phenomena that pertains to the process of creolization.

Against the thesis of creolization, Castilho (2007) gives a different explanation for the loss of number agreement, claiming that “a ‘ampla variação de concordância que ocorre em solo brasileiro’ dá continuidade a propriedades do português europeu não padrão” (2007:13) – (‘the wide agreement variation occurring in the only BP gives continuity to properties of nonstandard EP’ – my translation). Together with Scherre (2007) and Scherre & Naro (2007), he claims that there is no structural difference between EP and BP. Furthermore, they reject the explanations based on African or indigenous influence, “dada a não concentração dos fatos examinados nas áreas povoadas por essas etnias” (12) – (‘since the facts under study do not concentrate in the areas inhabited by these ethnic groups’ – my translation). Naro &

---

<sup>88</sup> Helvécia Creole Portuguese is defined by Holm as a Portuguese-based creole spoken in an isolated village in the southern part of Bahia, a state of Brazil. This creole was discovered in 1961. Helvécia was founded by Swiss and Germans. This explains the name Helvécia, which is the Latin name for Switzerland (Holm 1992:45f.).

Scherre further assume that the loss of the plural marker *-s* in BP was rooted in EP dialectal form, which, in turn, had continued a preromantic drift (thesis of ancientness) - (Naro & Scherre 2007:36). The “European drift” met in Brazil other inputs such as the pidginization, which had a certain influence over Portuguese through the ‘Língua Geral’ (lit. ‘general language’), a variety of Tupi, and the European ‘língua de preto’. So, Portuguese was learned as L2. The authors do not completely exclude the possibility for a creole to have existed, but if so, it disappeared without leaving any traces in the documentation (47). The following sections will focus on the nonstandard variety, often referred to as ‘spoken Brazilian’ in the literature.

#### 5.4.1 The determiner system of Brazilian Portuguese

As in EP, articles in BP, both definite and indefinite, inflect for gender and number. This differs greatly from any creole language that we have analyzed so far, which do not have morphological gender distinction and their articles do not inflect for number either. Recall that articles in BP are the same as in EP: *um, -a* ‘a/an’ and *uns, umas* ‘some’, are the inflected forms of the indefinite determiner for singular (masculine and feminine) and plural (masculine and feminine), respectively. On the other hand, the definite determiner has the forms *o, a* ‘the’ for the singular, masculine and feminine. The plural is *os, as* ‘the’ also inflected for masculine and feminine.

As for definites, Kester & Schmitt (2007) assume them to be used in the following cases: unique reference, anaphoric (associated with an antecedent in the discourse via identity or a part-whole relation), and referential and attributive uses (117). These assumptions find practical evidence in the following sentences from Kester & Schmitt (2007):

- 356) *Esse/\*O copo está vazio e este/\*o copo está cheio.*  
‘This/ The glass is empty and this/the glass is full.’
- 357) *Eu comprei um bolo. O bolo receberu muitos elogios.*  
‘I bought a cake. The cake received many compliments.’
- 358) *Eu comprei um bolo. A cobertura não era muito boa.*  
‘I bought a cake. The frosting was not very good.’



The ungrammaticality of the definite determiner in (356) confirms the claim made above about the case of unique reference. In (356), indeed, there is no reference to a unique entity, which is why the demonstrative, and not the definite, is felicitous in such a context.

In the second clause of (357), the only possibility for *bolo* ‘cake’ to be felicitous in this context, namely anaphora to the identical antecedent *bolo* in the first clause, is to be introduced by the definite determiner. A bare noun would not be allowed in this position. The same holds in (358), where *cobertura* ‘frosting’ in the second sentence stays in a part-whole relation to its antecedent *bolo* in the first sentence. Here again, the anaphoric *cobertura* has to be introduced by the definite article. In fact, a bare noun would not be allowed.

Furthermore, a NP introduced by a definite determiner could be (non)referential or yield a kind reading. Take a look at the sentences below:

- 359) Estou procurando o decano.  
‘I am looking for the dean.’
- 360) A baleia é um animal mamífero.  
‘The whale is a mammal.’

In (359) *o decano* ‘the dean’ could be referential, namely referring to a certain (known) dean, or nonreferential (‘attributive’, in Kester & Schmitt 2007:117-119). Again, a bare noun would not be felicitous here. In (360) *a baleia* ‘the whale’ yields a kind reading, denoting the entire set (kind) of whales, and not to a specific member of this class. A kind reading in BP may be also yielded by a bare (singular) noun.

Exactly as it happens in English, the definite determiner in BP is also used for the first mention of the noun, namely when the noun was not already mentioned in the discourse, at least in contexts of kind-denoting expressions or kind consisting of a unique member, as shown in sentences (361-362) below (Kester & Schmitt 2007:120). Such a use of the definite determiner has been named “expletive use” by Vergnaud & Zubizarreta (1992). In BP, as remarked by Kester & Schmitt, we find another case of expletive use of the definite determiner, namely when definites precede the proper name (e.g. *A Maria* ‘(the) Maria’), as in Florentine and Northern Italian Romance varieties. This is called “non-referential use of the definite” (2007:121).

- 361) Ninguém sabe quem inventou a roda/\*roda.  
‘Nobody knows who invented the wheel/\*wheel.’

- 362) O sol/\*Sol está queimando sem misericórdia.  
'The sun/\*Sun is burning without mercy.'

A latter case where we find the definite determiner is the inalienable construction, as shown in (363) below:

- 363) Ele lavou o rosto.  
He washed the face  
'He washed his face.'

As for indefinites, "the indefinite singular can have wide and narrow scope, can be used in generic constructions and can be used to disambiguate between singular and plural readings" (Kester & Schmitt 2007:122). The uses of (singular) indefinites in BP described so far are illustrated in the sentences below. The indefinite NP *um brasileiro* in (364) may yield (non)specific readings, whereas the indefinite NP *um hulandês* in (365) receives a kind reading.

- 364) Maria quer casar com um brasileiro.  
Maria want.3sg marry with a Brazilian  
'Maria wants to marry a Brazilian.'
- 365) Um hulandês fala holandês.  
A Dutch speak.3sg Dutch  
'A Dutchman speaks Dutch.'

Before concluding this section, it is worthwhile to briefly look at the plural formation in BP. The process is actually the same as in EP and many of the creoles analyzed so far, but as we said at the beginning of the present section, one of the most striking characteristics of BP with respect to EP is the loss of number agreement within the NP. So, when a noun is introduced by a determiner, which is overtly marked for plurality, the noun often lacks the plural marker. Usually, as explained in Holm (1992), (P)BP marks plurality on the first element of the NP only, which is typically a determiner. For the other elements of the NP, plural marking is an option. In other cases, it is not the first element to be marked as plural. Take a look at the following examples, both from PBP (Holm 1992:52).

- 366) Um dos mais velho orixás.  
'One of the most ancient deities.'
- 367) O meus irmão.  
'(the) my siblings.'

In the educated dialects, the plural may be marked also on the noun and on further elements eventually present in the NP. This is shown by the following example from Kester & Schmitt (2007):

- 368) As casas grandes.  
The-pl house-pl big-pl  
'The big houses.'

So far, we have described a situation that is quite different from EP, where both the noun and its determiner (and eventual modifiers) are overtly marked for plurality.

#### 5.4.2 Bare nouns in Brazilian Portuguese

Brazilian Portuguese shows a widespread use of bare (count) nouns, but if compared to creoles, their distribution is more restricted. It depends on the fact that the use of overt determiners in BP is more similar to that of EP, albeit more restricted.

According to Schmitt & Munn (2001), BP generally allows bare nouns in argument positions. Bare singular count nouns, bare plurals, and bare mass nouns. Following Schmitt & Munn, Kester & Schmitt (2007) analyze bare singulars in BP as names of kinds and assume that they behave like English bare plurals. This assumption is based on the following cues: first, they may occur as subjects of predicates that apply to kinds only; second, they may yield generic and existential readings, depending on the predicate that accompanies them, and last, whenever elements such as negation, intensional verbs and durative adverbials occur in the same sentence as the bare singular, it allows narrow scope reading only (2007:110-112). These facts are shown in the sentences below from Kester & Schmitt (2007):

- 369) Cabrito é muito comum em Curaçao.  
Goat is very common in Curaçao  
'Goats are very common in Curaçao.'
- 370) Eu adoro gato.  
Isg love cat-pl  
'I love cats.'
- 371) Eu cumprei casa.  
Isg buy.PAST.Isg house  
'I bought a house/houses.'

- 372) Eu não vi                   mancha no chão.  
 1sg Neg see.PAST.1sg spot    on+DEF floor  
 ‘I didn’t see spots on the floor.’

The sentences in (369) and (370) yield generic readings for the subject and the object, respectively. More precisely, *cabrito* in (369) is the subject of the predicate ‘to be common’, which typically yields a kind reading interpretation. The same holds true for (370), which is a generic sentence, since it describes the habitual property of the speaking subject ‘I’ to love cats. On the other hand, the objects in the sentences in (371) and (372) receive existential readings. As assumed above, the case in (371) is easily explained by the fact that the predicate *cumprar* ‘buy’ and its past perfective reading typically yield existential readings. Finally, in (372) it is the negation that yields the existential (narrow scope) reading.

As for bare count nouns in the subject position, BP allows bare singulars as subjects of episodic sentences, as in the example below, where both bare singular count nouns receive a plural interpretation:

- 373) Mulher tá lendo    e    homem tá escrevendo.  
 Woman is reading and man    is writing  
 ‘Women are reading and men are writing.’

However, the following tendencies have to be taken into account. The bare singular can occur as the subject of individual-level predicates, but not as the subject of stage-level kind predicates, where a definite NP is required. In other words, bare singulars in the subject position have a restricted distribution. To better understand this, take a look at the opposition between the two sentences below (Kester and Schmitt 2007). Only sentence (374) allows a bare singular as subject since the predicate is at the individual level and not at the stage level as in (374), where a definite NP is required.

- 374) O dinossauro/\*dinossauro está extinto.  
 ‘The dinosaur/dinosaur is extinct.’  
 375) Dinossauro é (um) animal extinto.  
 ‘Dinosaur is an extinct species.’

Schmitt & Munn (2003) claim that bare singulars in BP are underspecified for number. A first consequence is that bare singular count nouns may yield both singular and generic readings (but not mass ones – Kester & Schmitt 2007). Further evidence is that “bare singulars force durative aspect on predicates that are sensitive to whether their objects are quantized or not”

and [...] “if bare singulars were specified for number, they should force a terminative reading of the VP, but they do not”. This is shown in the example below from Munn & Schmitt (2001).

- 376) Eu escrevi carta por muitos anos/#em uma hora.  
I wrote letter for many years/ in one hour.  
'I wrote letters for many years/in one hour.'

As for the object position, the bare singular usually receives a plural reading whenever a time-span aspectual adverbial is present or when the BN occurs as the object of a non-atomicity entailment (2007:127), as illustrated in (377) and (378), respectively:

- 377) Eu matei iguana por duas horas.  
1sg kill.1sg.PAST iguana for two hours  
'I killed iguanas for two hours.'
- 378) Pedro colecciona selo.  
Pedro collect.3sg stamp  
'Pedro collects stamps.'

Following Kester & Schmitt (2007), discourse anaphora provides further evidence for the unmarkedness of bare singulars as for number. In fact, the antecedent of a plural pronoun could be a (morphologically) singular bare noun. This does not hold in the opposite case, namely if the antecedent is a bare plural, then the pronoun must be plural as well.

- 379) Eu tenho filho. Você quer conhecer eles/ele?  
I have child. You want meet them/him  
'I have children/a child. Do you want to meet them/him?'

At this point, we should note that singular BNs in BP yield singular readings, for example with verbs that entail atomicity (e.g. ‘marry’, at least in monogamous societies, as remarked in Kester and Schmitt 2007:127):

- 380) Maria quer casar com brasileiro.  
Maria want.3sg marry with Brazilian  
'Maria wants to marry a Brazilian.'

Furthermore, following Kester & Schmitt (2007), bare singulars are not felicitous in contexts of inalienable possession and with unique and familiar entities. To better understand this, take a look at (381) and (383) below. They are compared to their grammatically acceptable

counterparts in (363) and (362) above, re-proposed here in (381) and (383). *Rosto* ‘face’ in (382) and *sol* ‘sun’ in (384) must be preceded by a definite determiner in order to be grammatically acceptable.

- 381) \*Ele lavou rosto.  
 3sg wash.PAST.3sg face  
 ‘\*He washed face.’
- 382) Ele lavou o rosto.  
 ‘He washed his face.’
- 383) \*Sol está queimando sem misericórdia.  
 sun is burning without mercy  
 ‘The sun is burning without mercy.’
- 384) O sol está queimando sem misericórdia.  
 DEF sun is burning without mercy  
 ‘The sun is burning without mercy.’

As for bare plurals in BP, they behave like English bare plurals. They may occur in both subject and object positions, and may yield both generic and episodic readings (Schmitt & Munn 2003). In the subject position, they seem to be perfectly felicitous in both generic (385) and episodic contexts (386):

- 385) Cabritos são muito comuns em Curaçao.  
 Goat-pl are very common in Curaçao  
 ‘Goats are very common in Curaçao.’
- 386) Tem computadores na minha mesa.  
 Have computer-pl in+DEF my table  
 ‘There are computers on my desk.’

Importantly, BP bare plural subjects in episodic contexts (not existentials, hence different from (386)), are more felicitous if accompanied by a definite article, such as in EP (Kester & Schmitt 2007). The bare plural *alunos* in (387) is a bit odd, and a definite NP would be preferred.

- 387) Depois que ?alunos reclamaram, eles consertaram o filme.  
 After that ?students complained, they fixed the film.  
 ‘After students complained, they fixed the film.’

As said above, bare plurals are also allowed in the object position. Take a look at the following sentence (Kester & Schmitt 2007):

- 388) Eu vi livros na mesa.  
I saw book-PL on-the table  
'I saw books on the table.'

In sum, Kester and Schmitt (2007) conclude that BP displays the following nominal form-interpretation combinations: bare nouns and NPs with overt definite D may yield kind readings or can be interpreted as characterizing predicates. This latter reading, i.e. the characterizing one, is also available for singular indefinites. Definite DPs may also derive familiar and unique entity readings (not previously mentioned) and inalienable possession (this latter one seems to be available for possessives either) – (2007:123, table 2). In BP, the definite determiner is necessarily present in constructions of inalienable possession, with proper names and with certain kind-denoting expressions (130).

Concluding this present section on bare nouns in the argument position, “Brazilian Portuguese is like English in allowing empty determiners, but is also like the other Romance languages in allowing interpretable number to be absent when not otherwise required” (Schmitt & Munn 2003:13).

Now, it would be interesting to briefly look at bare nouns in the predicate position. According to Schmitt and Munn (1999, 2003), in this regard, BP behaves like the Romance languages and differently from English. Indeed, bare nouns in BP can quite freely occur in postcopular positions. As pointed out by the authors, there are some other positions where bare singulars are quite freely allowed: *reprise-commentaire*, *as*-construction, and part-whole *with* constructions. Take a look at the examples below from Schmitt and Munn (1999):

- 389) Pedro trouxe uma mesa de camping, mesa de camping que serviu para o picnic.  
'Pedro brought his camping table, a camping table that served for the picnic.'
- 390) Ninguém poderá usar-nos/usá-lo como testemunha.  
Nobody will be able to use us/him as witness.  
'Nobody will be able to use us/him as witnesses/a witness.'
- 391) Comprei um cachorro de/com rabo comprido.  
I bought a dog of/with long tail  
'I bought a dog with a long tail.'

The reprise-commentaire is exemplified in (389) and is also found in other Romance languages like French, Italian and Spanish. In English, on the other hand, it is not possible, at least not with a bare noun. Indeed, an indefinite determiner has to be present in order to allow such construction. In Schmitt & Munn's words, "the reprise-commentaire functions as a secondary predicate on the noun phrase and as such clearly confirms the prediction that predicates can be bare in Romance" (1999:7).

In (390), we find an example of *as*-construction, which is expressed via *como* 'as, like' in BP. This is another predicative construction, and here again in BP a bare noun is allowed, whereas English requires the introduction of an indefinite article.

Finally, in (391) the part-whole *with* construction, expressed in BP via *de* or *com*, the bare singular noun is licensed. Again, in English an indefinite determiner has to introduce the noun. This construction, as explained in Schmitt and Munn, functions as a modifier of the noun phrase.

Concluding, bare nouns in BP behave similarly to both English and the Romance languages, but from different perspectives. Bare count nouns in argument positions resemble the behavior of English bare plurals, whereas BP bare predicates pattern together with bare nouns in the predicate position in the other Romance languages.

## 5.5 Summary and comparison: BNs among creoles

There seems to be both commonalities and differences among the languages analyzed so far. For many aspects, Cape Verdean Creole, Santome, Papiamentu and Brazilian Portuguese pattern on a par with Guinea-Bissau Creole. In this section, we will briefly review the main points described so far with regard to the syntax and semantics of bare noun phrases of these languages, thus highlighting both similarities and differences.

There are important commonalities among the languages analyzed so far; however, there are also some crucial differences to take into account. The semantics of number marking seems to be similar in most of the languages under consideration, but for Brazilian Portuguese. Overt plurality, indeed, combines with specificity in GBC, CVC, Santome, and Papiamentu. Animacy, on the other hand, does not seem to play any important role in the



plural marking in Papiamentu; however, in the cases of GBC, CVC, and Santome, animacy is crucial as well as specificity (see Table 3).

As for the distribution of bare nouns, the situation seems to be quite homogeneous: all languages involved in the present comparison allow both bare singulars and bare plurals in both the subject and object positions, but for Papiamentu. As we can see in Table 4, in fact, Papiamentu allows a more restricted distribution of bare plurals in both the subject and object positions. More specifically, bare plural subjects are allowed only in episodic sentences, whereas bare plural objects occur only if modified and only in episodic sentences.

As for the semantic behavior of bare nouns, our findings show again a high degree of similarity among the languages taken into account here. All languages allow for their bare nouns, both singular and plural, generic and existential readings, (in)definite and (non)specific interpretations. There are, however, some exceptions to this common tendency: first, Santome does not allow a nonspecific reading of singular bare nouns; second, in Papiamentu bare plurals cannot yield generic (nonspecific) readings; finally, bare singulars in BP cannot derive a generic reading. These findings are summarized in Table 5. As for number interpretation, Table 6 summarizes the facts. The tables are inspired by and partially based on Baptista (2007b).

Table 3. *Plural markers, animacy, and referentiality*

	Animacy	Referentiality
GBC	Yes	Yes
CVC	Yes	Yes
Santome	Yes	Yes
Papiamentu	No	Yes
BP	N.R.	N.R.

N.R. means that the features, namely animacy and referentiality, do not seem to be relevant for overt plurality in BP.

Table 4. *Syntactic position of bare singulars and plurals*

	Subject: Bare SG	Subject: Bare PL	Object: Bare SG	Object: Bare PL
GBC	Yes	Yes	Yes	Yes
CVC	yes	yes	yes	yes
Santome	Yes	No	Yes	No
Papiamentu	Yes	Yes/No	Yes	Yes/No
BP	Yes	Yes	Yes	Yes

Note: yes/no in Papiamentu as for subject and object bare plurals means that bare plural subjects are possible only in episodic sentences, whereas bare plural objects occur only if modified and only in episodic sentences.

Table 5. *Semantic interpretation of bare nouns*

	Generic	Exist.	Definite	Indef.	Specific	Nonspec.
GBC	Yes	Yes	Yes	Yes	Yes	Yes
CVC	Yes	Yes	Yes	Yes	Yes	Yes
Santome	Yes	Yes	Yes	Yes	Yes	Yes/No
Papiamentu	Yes/No	Yes	Yes	Yes	Yes	Yes/No
BP	Yes/No	Yes	Yes	Yes	Yes	Yes

Notes:

Santome: the nonspecific reading is not available for singular BNs.

Papiamentu: the generic reading cannot be yielded by bare plurals. The nonspecific reading is not available for bare plurals.

BP: bare singulars cannot yield a generic reading.

Table 6. *Number interpretation of singular bare count nouns*

	Singular	Plural
GBC	Yes	Yes
CVC	Yes	Yes
Santome	Yes	Yes
Papiamentu	Yes	Yes
BP	Yes	Yes

## 5.6 Mandarin Chinese

Before we start with the description of the nominal system of Mandarin Chinese and its bare nouns, let me give some information about Mandarin. First of all, we should clarify our use of the name ‘Mandarin’ because it is, in fact, somewhat ambiguous. According to Li and Thompson (1975), it indicates either the standard and official language of the mainland and Taiwan, or designates “the natural variety of Chinese otherwise referred to as “the Northern-Chinese dialect”” (259). In this latter meaning of the name, i.e. as a natural language, Mandarin is spoken on a very vast area, namely north of the Yangtze River and in the southern provinces of Yunnan and Guizhou (ibid.). Still, the internal variation is not very high, and there seems to be a high degree of mutual intelligibility among Mandarin varieties. Mandarin and, more generally, Chinese languages are typologically quite different from the Indo-European language family. One of the most striking differences is represented by their nominal systems. We will see this in more detail below, but let me anticipate that Mandarin (as well as Cantonese) lack articles and number morphology. Therefore, Mandarin must find another way for expressing semantic categories such as definiteness and indefiniteness, which, on the other hand, find realizations in many Indo-European languages in the use of determiners. As we will see later, Mandarin expresses (in)definiteness via bare nouns or expressions made of classifier + noun. Classifiers are also used for expressing number, both singular and plural, whenever it needs to be overtly specified. Otherwise, we simply find bare nouns.

Before we turn to the Chinese nominal system and see in more detail what has been introduced so far, let me address the linguistic situation of China, where different other languages (often referred to as dialects) are spoken. Following Li and Thompson (1989), the following seven are the principal dialect groups in China: Mandarin (Northern, Northwestern, Southwestern, and Lower Yangzi), spoken by about 70 % of population, Wú, Xiāng, Gàn, Hakka, Mǐn, and Yuè.<sup>89</sup> The Chinese language family is classified as an independent Sino-Tibetan language family.<sup>90</sup>

### 5.6.1 On word order in Mandarin

It seems that in many languages the natural order for definiteness is preverbal, whereas indefiniteness finds realization more likely in postverbal positions. The (in)definite status of a noun seems to take precedence over its subject/object role. In Mandarin Chinese, “in the absence of morphological markers, word order has taken on the function of denoting the definite/indefinite property of nominal” (Li and Thompson 1975:165). As evidence of this fact, definite nouns, independent of their syntactic position as arguments (subject or object), are usually placed in the preverbal position. In the case of objects, a preposition *ba* is used.<sup>91</sup> On the other hand, indefinites are tendentially postverbal. In the case of subjects, they may precede the main verb but are preceded by an auxiliary predicate *you* ‘exist’. Moreover, Li and Thompson claim, with respect to word order in Mandarin Chinese, that this kind of word order was developed in the past millennium and that it conflicts with the shift from SVO to SOV, which is claimed to be a diachronic process.

Word order in Mandarin is in fact not always of the type SV, but the subject may also be postverbal (VS), principally if the subject of an intransitive predicate is not morphologically marked for definiteness. Hence, if preverbal, the subject will yield a definite interpretation, whereas a postverbal subject will be interpreted as indefinite.

---

<sup>89</sup> Along with Mandarin, most of the other language groups of China are further internally subdivided. For more details, see Li and Thompson (1989).

<sup>90</sup> In the Sino-Tibetan language family, the other major branches are: the Tibetan languages (languages of Tibet), Lolo-Burmese (principally Burma; also Southern China, Southeast Asia, and the Tibetan borderland) and Karen (lower Burma and the southern border region between Burma and Thailand); (Li and Thompson 1989:2).

<sup>91</sup> In case of *ba*-constructions (usually described as indicating ‘affectedness’), the post-*ba* NP, i.e. the internal argument of the predicate, can be the semantic subject of the sentence. Moreover, whenever *ba* is used, the verb does not show up in its usual position, namely between the subject and object, but at the end of the sentence (Huang, Li and Li 2009).

In sum, Mandarin is a language with an SVO word order, but is in the process of changing to SOV.

### 5.6.2 Chinese nominal system: Count vs. mass distinction

In the preceding section, we introduced some typological issues concerning Mandarin: we said that its nominal system does not share some important elements such as determiners and number morphology (at least, not directly on the noun via suffixes). Moreover, Mandarin is a topic-prominent language: it means that the topic has crucial importance in Mandarin sentences. As a further crucial difference from Indo-European languages, or at least from most Romance and Germanic languages,<sup>92</sup> bare nouns occur quite freely in argument positions in Chinese languages. This seems to resemble the situation described above as for the creoles (and noncreoles) taken into account in the present study (GBC, CVC, Santome, Papiamentu, and BP as well). However, there are also important differences, which we will see in the discussion below.

Turning to the case of Mandarin, we already introduced the fact that they lack determiners altogether. Determiners are usually the locus for expressing semantic categories such as definiteness and indefiniteness. In many languages, there are also different morphological forms for determiners to express number, such as German, Italian or Portuguese, but not in English (at least for definiteness). If Mandarin does not have determiners, how can it express such semantic categories? Recall that one of the proposals of the present study is to find evidence for the hypothesis that every language somehow (overtly or covertly) realizes these categories. Mandarin does it via bare nouns. As for number, it uses classifiers or simply bare nouns. Another striking difference between Mandarin and Indo-European languages is that Mandarin has classifiers, which are morphological items that express a measure, so to say, for the noun to be countable. If not, nouns are mass, and “they are called ‘classifier’ because different nouns have different count-classifiers, depending on the shape or any property of the individual units that come with the natural partitioning” (273): e.g. *ge* is a classifier for “long, tall things like humans”. This classifier goes back to a word meaning “bamboo” (Cheng and Sybesma 2005:273). According to Krifka (1995),

---

<sup>92</sup> The need for specifying that it is the case of Romance and Germanic comes from the fact that other Indoeuropean languages have a much more widespread use of bare nouns. For instance, Russian lacks determiners altogether, both definite and indefinite. It has, however, number morphology, thus resulting different from Chinese languages with respect to the nominal system.

common noun constructions in Mandarin are more transparent than in languages such as English. Mandarin can resort, indeed, to classifiers for specifying if a noun is being used as generic (taxonomic) or object-referring. The following examples from Krifka (1995:398f.) will exemplify these facts:

- 392) a. Sān zhī xiong  
three CL bear  
'three bears' (object reading)
- b. Sān zhǒng xiong  
three CL bear  
'three bears' (species; taxonomic reading)

To better understand the function of classifiers, we should look at the distinction between mass and count in Mandarin. This distinction is reflected in the classifier system, hence Mandarin has both count and mass classifiers. Against Chierchia (1998), Cheng and Sybesma claim that not all nouns in Mandarin are mass, although they seem so, since Mandarin nouns do not have number morphology. This is true for both mass and count nouns. Cheng and Sybesma assume that Chinese has a distinction between mass and count nouns, just as English does. Mass nouns like 'water' need a unit to be counted. Hence, measure words make mass nouns countable. Count nouns are countable for their own nature or, rather, they "have a built-in semantic partitioning" (Cheng and Sybesma 2005:273). So, measure words create a unit to count by and, on the other hand, classifiers "simply *name* the unit that the semantic representation of the noun naturally provides" (ibid.). In Chinese, both mass and count nouns need a counter. According to Cheng and Sybesma, "in the case of mass nouns, the counters (the measure words, or mass-classifiers) create their unit of counting, in the case of count nouns the counters (the classifiers, or count-classifiers) simply name the unit that the semantic representation of the noun provides" (ibid.). Classifiers do not have only a counting function, but also an individualizing one. Canonical classifiers are always singular. Hence, we may say that "the classifier singles out one entity from the plurality of entities provided by the semantic representation of the noun in the lexicon; it picks out one instance of what is denoted by N" (276).

As for plurality, we already said that Mandarin lacks number morphology. Plurality finds overt realization via the classifier *xie*, which is the classifier for the unspecified plural. However, Mandarin generally expresses plurality using the numeral, and not the plural classifier. If we want to say 'three books', we use plural overt morphology in English. As a

crucial difference, Mandarin does not use the plural, but the canonical classifier preceded by the numeral (Cheng and Sybesma 2005): *San ben shu* = three CL<sup>VOLUME</sup> book ('three books').

In present-day Mandarin, there are non-definite articles, but demonstratives: *zhèi* 'this' and *nèi* 'that'. Recall that the tendency claimed so far, namely that the postverbal position is associated with indefiniteness, is true so long as the noun is not marked for definiteness. On the other hand, a noun introduced by a demonstrative can be postverbal with transitive verbs. The same holds for proper names and pronouns, which are said to be inherently definite. The following sentences will show these facts:

- 393) *Háizi dǎ-pò le nèige chuānghu.*  
Child hit-be asp broken that window  
'The child broke that window.'
- 394) *Tā dǎ le Zhāng-san / wǒ le.*  
He beat asp Zhang-san/me sfp  
He has beaten Zhang-san/me.

According to Li and Thompson, there are also cases where nouns, neither marked nor inherently definite, are found in the postverbal position:

- 395) *Qǐng nǐ guān mén.*  
Please you close door  
'Please close the door.'

What about anaphoric nouns, namely nouns which have been already referred to in the discourse context? Interestingly, Li and Thompson note that an anaphoric noun cannot be in the same position as a specific noun like in (395), where *mén* is in the postverbal position, but interpreted as definite on pragmatic basis. If we want to answer the question "what about the key(s)?", then the noun 'key' cannot be bare and in the postverbal position. It must be either left-dislocated (focalization), should be introduced by the object marker *ba*, or simply omitted.

- 396) *Yàoshi, wǒ wàng le / wǒ bǎ yàoshi wàng le / wǒ wàng le.*  
Key, I forget asp / I obj.marker key forget asp / I forget asp

As for the preverbal position (often, sentence initial), the noun is always interpreted as definite, even if introduced by the numeral *yī* 'one', which is not an indefinite determiner, since Mandarin is said to lack determiners altogether. Thus, *yī* is a classifier (singularizer),

here acting like a quantifier. It seems that nouns marked by the classifier *yi* are not only singularized, but also generic. A specific interpretation does not seem to be ungrammatical, but a generic one sounds more natural, following Li and Thompson. This seems to contradict what we said so far with respect to the definite interpretation of preverbal nouns. Nonetheless, the authors explain these facts with the assumption that genericity implies somehow definiteness, because “generic noun phrases should be considered a type of known information because they are established in the permanent registry of discourse” (175).

- 397) Yige rén           shuì-jiào-de shíhòu,      chángcháng zuò mèng.  
 One.CL-person sleep subord.marker time, often make dream  
 ‘When a person sleeps, he often dreams.’

However, there is a way to have an indefinite subject. It must be preceded by the existential verb *youǐ* ‘exist’. Li and Thompson explain this case by claiming that the indefinite is however postverbal, in fact *youǐ* is a verb. Clearly, a noun phrase introduced by *youǐ* will yield an existential reading:<sup>93</sup>

- 398) Youǐ yige rén   dǎ-pò nège chuānghu le.  
 Exist one person hit-be broken that window asp  
 ‘A person broke that window.’

Before concluding this section on Mandarin, let me briefly address PPs.

So far, we have principally dealt with nouns in the subject or object position, but there is something interesting to note also in cases of prepositions. More specifically, in passive constructions, the agent marker *bèi* may yield both definite and indefinite readings of the noun it introduces. This seems to contradict what we said so far with respect to interpretation tendencies in Mandarin. Li and Thompson (1989) explain this fact as follows: passive constructions in Chinese are under Western influence, mostly English. In fact, under Western influence the use of passives in Mandarin has increased, but until the beginning of the 20<sup>th</sup> century, the passive was used only with verbs with negative meanings such as ‘beat’. This

---

<sup>93</sup> According to Huang (1987), the construction with the verb *you* is one of the ways Chinese has in order to build an existential sentence. Unlike Li and Thompson, Huang glosses *you* as ‘have’ and not as ‘exist’: *you gui*, lit. ‘have ghost’ (‘there are ghosts (here)’) (227). The other two possibilities for the existential sentence, as described in Huang, are, on the one hand, sentences with verbs such as *lai*, ‘come’, *fasheng*, ‘happen’ or *dao*, ‘arrive’, and, on the other hand, with locational verbs, “transitive or intransitive verbs that subcategorize for a locative phrase” (228). Examples of these two latter ways are: *fasheng le yijian chehuo*, happen Perf one accident (‘an accident happened’), and *chuang-shang tang-zhe yige bingren*, bed-top lie-Dur one patient (‘in the bed lies a patient’) – (Huang 1987:228).



tendency has probably brought with itself another tendency, namely the nonrestriction on the definiteness of the agent (179f.).

### 5.6.3 Bare Nouns in Mandarin

We have already introduced the fact that bare nouns in Chinese languages are very widespread. They may occur in both the preverbal and postverbal positions, namely in subject and object positions, respectively, since Mandarin is an SVO language.

Also in Cantonese bare nouns are widespread, but Cantonese makes more use of classifiers than Mandarin. In fact, Cantonese uses classifiers for expressing (in)definiteness, whereas (in)definiteness in Mandarin relies on bare nouns. This implies that the occurrence of bare nouns in Cantonese is lower than in Mandarin, which is why I will not include Cantonese in this description of Chinese bare nouns.

Bare nouns in Mandarin show a wide array of interpretations. According to Cheng and Sybesma (1999), as for BN interpretation in Mandarin, postverbal BNs may be interpreted as definite, indefinite, and generic. The following examples from Cheng and Sybesma (1999) will show these facts. In (399-401) we have bare nouns in the object position; *shu* ‘book’ in (399) yields an indefinite reading, singular or plural; *tang* ‘soup’ in (400) receives a singular definite interpretation, and finally *gou* ‘dog’ in (401) is interpreted as plural generic.

- 399) Hufei mai shu qu le.  
Hufei buy book go SFP<sup>94</sup>  
‘Hufei went to buy a book/ books.’
- 400) Hufei he-wan-le tang.  
Hufei drink-finish-LE soup  
‘Hufei finished the soup.’
- 401) Wo xihuan gou.  
I like dog  
‘I like dogs.’

Unlike Li and Thompson, who put the stress on the position of nouns for their key interpretation, Cheng and Sybesma (2005) claim that the interpretation of a bare noun in

---

<sup>94</sup> *Le* has (at least) two functions in Mandarin grammar. It is a sentence-final particle (SFP), and it also lexicalizes aspect (Cheng and Sybesma 1999:510, Li and Thompson 1989 and Krifka 1995 among others).

Mandarin depends on the predicate with which it occurs. A bare noun in Mandarin may be interpreted as indefinite, definite or generic. With unbounded activity verbs (e.g. ‘go to buy’ in sentence (399)) the BN yields an indefinite reading. On the other hand, with bounded events (‘finish the soup’, (400)), the BN gets a definite reading. Finally, with unbounded states (401), the BN is generic. Mandarin bare nouns in the preverbal position may yield definite or generic readings. Hence, as a crucial difference from BNs in the object position, subject BNs cannot yield an indefinite reading. The sentences in (402-404) below (Cheng and Sybesma 1999) will show this:

- 402) Gou yao guo malu.  
 dog want cross road  
 ‘The dog wants to cross the road.’
- 403) Gou jintian tebie tinghua.  
 dog today very obedient  
 ‘The dog/ dogs was/ were very obedient today.’
- 404) Gou ai chi rou.  
 dog love eat meat  
 ‘Dogs love to eat meat.’

The subject bare noun *gou* ‘dog’ in (402) yields a definite (referential) interpretation. An indefinite interpretation is excluded by the authors. The definite reading is probably conveyed by the intensional verb in the sentence, and the fact that the sentence is episodic. As a difference, in (404) the same bare subject receives a generic interpretation. The predicate ‘to love’ is in fact an individual-level predicate, which tendentially accompanies generic statements. Number interpretation is plural because of the genericity of the whole sentence, and not because more instances of the kind dog are referred to (in fact, this is not an object reading). Moreover, it cannot be interpreted as singular since in that case it would correspond to a taxonomic reading (Krifka et al. 1995), as we already saw in Chapter 2 of the present study. Finally, the sentence in (403) is an episodic statement, and the bare subject *gou* may be both singular and plural (presumably, the correct interpretation depends here upon the context).

Table 7 may help us summarize the above described facts about the distribution and interpretation of bare nouns in Mandarin. The following table is adapted from Cheng and Sybesma (1999:512).

Table 7. *Distribution and interpretation of BNPs vs. CLPs in Mandarin*

	Indef	Def	Generic
BN	Yes	Yes	Yes
CL + N	Yes	No	No

As for classifiers, we already said that classifiers are singularizers. This means that they behave like D, say, in English in some respects, and differently in other respects (520f.). More specifically, the definite determiner ‘the’ in English may introduce both singular and plural nouns, depending on whether they are marked for singularity or plurality. In Mandarin, on the other hand, there is no plural marker, so Chinese lacks any nominal number morphology.

As for definiteness in Mandarin, this is expressed via bare nouns, and not via classifiers, as in the case of Cantonese. Moreover, recall that definite (bare) nouns in Mandarin have a freer distribution than their indefinite counterparts. In fact, indefinite bare nouns are restricted to internal argument positions, whereas definite nouns may occur in both subject and object positions.

Turning to classifiers in Mandarin, Cl+N – phrases are restricted to nonspecific indefinites. They cannot get a generic reading, either. Furthermore, Cl+N - phrases cannot occur in the preverbal position. Thus, according to Cheng and Sybesma (2005), they occur only as objects of unbounded activity predicates:

- 405) Wo xiang mai ben shu.  
 I want buy CL<sup>VOLUME</sup> book  
 ‘I would like to buy a book.’<sup>95</sup>

Following Cheng and Sybesma (2005), Num+Cl+N-phrases can occur as objects of bounded predicates, and can be interpreted as specific. The numeral *yi* followed by a classifier together with a noun may yield both specific and nonspecific readings. It may also mean ‘one’; however, both Cl+N and Num+Cl+N are indefinites and may occur in the postverbal position only. Num+Cl+N may be (non)specific, whereas Cl+N may only be nonspecific. *Yi-ben shu* in (406) may be both specific and nonspecific.

<sup>95</sup> Cl+P-phrases such as *ben shu* in (21) have often been described as phonologically reduced forms preceded by the numeral *yi* ‘one’. The full original form would thus be *yi-ben shu* (Cheng and Sybesma 2005:262). Cheng and Sybesma, however, do not agree with this explanation.

- 406) Wo xiang mai yi-ben shu.  
I want buy one-CL book  
'I would like to buy a book.'

In sum, definiteness and genericity in Mandarin may be expressed via bare nouns only. On the other hand, indefiniteness finds realization via bare nouns, CI+N and Num+CI+N.

## 5.7 Vietnamese

The official language of Vietnam, Vietnamese, is part of the Viet-Muong subgroup of the Mon-Khmer cluster, which belongs to the Austro-Asiatic language family. According to Nguyen (2004, 2008) it has over 80 million speakers in Asia, Australia, Europe, and North America. Vietnamese has three main dialects, namely the Northern, Central and Southern ones, and among these varieties there is a high degree of mutual understanding.

Similar to Chinese and Gbe languages, Vietnamese is a tone language that has six distinct tones. Tones in Vietnamese have a lexical function, more specifically they distinguish between word meanings. Following Nguyen (2004, 2008), like South-East Asian languages in general, Vietnamese is an SVO language with a rigid head-initial order. Moreover, it is isolating: "Vietnamese lacks overt markings of Case, Number, Gender, Tense or finite/non-finite verb status" (2008:8).

### 5.7.1 The nominal system of Vietnamese

As we mentioned above, Vietnamese is an isolating language. This means that it lacks overt markings of Case, Number, and Gender, as for the nominal system.

With respect to word order, there are both prenominal and postnominal items. In greater detail, Nguyen (2004, 2008) counts among the former ones, i.e. the prenominal modifiers, items such as quantifiers, articles, numerals, the particle *cái*, classifiers, and measure phrases. On the other hand, postnominal modifiers are noun adjuncts, adjective phrases, prepositional phrases, relative clauses, demonstratives, and possessives. We are going now to see these categories in more detail.

We already said that Vietnamese is a classifier language. According to Nguyen, classifiers always show up whenever a noun phrase contains a numeral. The main function of a classifier in Vietnamese is to individuate nouns and make them countable (2008:9). This is similar to the situation found in the discussion on Mandarin Chinese above. The sentences in (407) and (408) demonstrate an idea of Vietnamese classifiers:

407) Toy mua cuốn sách.

I buy CL book

‘I buy a/the book.’

408) Ba cây bút.

three CL pen

‘Three pens.’

The classifier *cuốn* in (407) introduces the noun *sách* ‘book’, which receives a singular reading, either definite or indefinite. Differently, the noun *bút* ‘pen’ in (408) receives a nonsingular reading. It depends on the presence of the numeral, which precedes the classifier. If the numeral were not there, then the noun would obligatorily receive a singular reading, since classifiers are singularizers.

Classifiers do not usually occur without the noun; however, whenever a modifier occurs, for example a demonstrative, an adjective or a relative clause, then the noun can be omitted and the classifier surfaces alone. In (409) we have an example, where a demonstrative shows up in its pronominal use (2008:11). We may already note that the demonstrative occupies the rightmost position of the nominal string, and we will see this in more detail below.

409) Cuốn (sách) mới.

CL book that

‘That one.’

Classifiers in Vietnamese are mostly used for referring to animate vs. inanimate. We saw above that it is an important distinction in creoles as well, at least with respect to overt plurality. According to Nguyen (2004, 2008), there is in Vietnamese a “three-way animacy-based distinction” (2008:11). There are two classifiers for animates, distinguishing between humans and nonhumans, i.e. *người* for humans (it literally means ‘person’), and *con* for

animals. As for inanimates, they are referred to via the classifier *cái*.<sup>96</sup> Crucially, such classifiers cannot co-occur in the same sentence (2004:12f.). Some examples are in (410-412) below:

- 410) Người bạn  
CL friend  
411) Con chó  
CL dog  
412) Cái bàn  
CL table

According to Nguyen (2004, 2008), there are several types of classifiers: kind-classifiers, event-classifiers, and unit-classifiers, which “provide countable units to individuate nouns” (2004:13), and those are what we saw above for the distinction animate vs. inanimate.

- 413) Hai loại chó.  
two CL:kind dog  
‘Two kinds of dogs.’  
414) Một cuộc họp.  
one CL:event meet  
‘A meeting.’

In the same position where classifiers occur, we may also find measure phrases, i.e. expressions indicating quantity. These latter ones are distinguished from classifiers on a syntactic basis<sup>97</sup> and, following Nguyen, can be of two types: conventional and nonconventional, or standard and nonstandard. To the former ones, there belongs measure expressions such as ‘kilogram’ and ‘meter’. On the other hand, noncanonical measure expressions are ‘a houseful (of)’, ‘a tableful (of)’, etc. The example in (415) is an instance of standard measure phrase, whereas in (416) we have a nonstandard measure expression:

---

<sup>96</sup> Nguyen (2008:12) reminds the reader to pay attention to syncretic forms. In fact, the classifier for inanimates, i.e. *cái*, is syncretic with the particle *cái*, with the adjective *cái*, which has the two meanings ‘main’ and ‘female’, and with the archaic noun *cái* ‘mother’. As for the classifier for animals, i.e. *con*, it is homonymous with the noun *con* ‘child(ren)’ and with the adjective *con* ‘small’.

<sup>97</sup> According to Nguyen (2008:21-31,116), measure expressions occur in the same surface position as classifiers, i.e. between the numeral and the noun. Nonetheless, they have a different syntactic structure. Measures would occur in phrases and are base-generated in a different syntactic position than classifiers, namely they are DPs base-generated in Spec,CLP.

- 415) Một kí đường.  
 one kilogram sugar  
 ‘a kilogram of sugar’
- 416) Một nhà đường.  
 one house sugar  
 ‘a houseful of sugar’

A further difference between classifiers and measure phrases is that the latter ones are a nonclosed group and, in principle, “any appropriate object can be used as a container to measure *đường* ‘sugar’” (Nguyen 2008:27).

Following Nguyen (2004, 2008), numerals in Vietnamese occur to the left of the classifier in the following order: Numeral + Classifier + Noun. The numeral always co-occurs with a classifier, and even in case of topicalization, whereby the noun is moved to the left periphery of the sentence, the numeral cannot be left alone. As we will see in more detail below, the classifier always remains close to the numeral.<sup>98</sup>

Quantifiers are a “closed lexical class of words” (57), such as *cả* or *tất cả*, both meaning ‘all’, and *mỗi* or *từng*, ‘each’. When a quantifier is required, it occurs in the first position of the nominal string (59):

- 417) Cả/tất cả ba mươi con trâu này.  
 all/all three ten CL buffalo this  
 ‘All these thirty buffalos.’

With certain quantifiers, and also with numerals, when the context is given and the referent is clear enough, the noun and its classifier may be omitted. The following examples from Nguyen (2008) will show this:

- 418) Vài (con tem).  
 a few CL stamp  
 ‘a few (stamps)’ answering to the question ‘how many stamps do you need?’
- 419) (Tôi muốn mua) ba (quả).  
 I want buy three CL  
 ‘(I want to buy) three.’ Answering to the question ‘How many oranges do you want to buy?’

---

<sup>98</sup> If the focus marker occurs in a [Num-CL-N]-phrase, it shows up between the number and the classifier: *ba cái con mèo này*, three CÀI CL cat this, ‘these very three cats’ (Nguyen 2008:57).

In (419) the numeral may occur alone, without the noun and the classifier.

As for postnominal modification of the noun, we already noted that NP adjuncts, adjectives, prepositional phrases, relative clauses, demonstratives and possessives belong to this group.

With respect to demonstratives, they occur as the last element in the nominal string. They do not undergo any morphological change as for number. In Vietnamese there is a three-way demonstrative system (2008:66): proximal (*này, này, ni*, ‘this, these’); medial (*đó, đấy, ấy, nó*, ‘that, those’); and distal (*kia, tê*, ‘that, over there’).

- 420) Ba con trâu trắng của tôi này.  
three CL buffalo white of I this  
‘These three white buffaloes of mine.’

The sentence above shows the rightmost position of the proximal demonstrative *này*. Furthermore, it also contains an adjective, which modifies the noun *trâu* ‘buffalo’, and a possessive. The adjective occupies the position at the immediate right of the noun and is followed by the possessive.

As for articles, it has often been claimed that Vietnamese does not have any, but Nguyen claims that it does.

Other examples of postnominal modification are found in the sentences below. The examples in (421), (422), and (423) contain a noun adjunct, a prepositional phrase and a relative clause, respectively.

- 421) Sách hình học.  
book geometry  
‘A/the geometry book.’
- 422) Căn nhà gần bờ sông  
CL house near bank river  
‘The house near the river bank.’
- 423) Cuốn từ điển mà tôi thích  
CL dictionary that I like  
‘The dictionary that I like.’

At the left of the classifier or, in case a numeral co-occurs, between numeral and classifier, a further item may occur, i.e. the so-called ‘extra *cái*’, which is analyzed by Nguyen as a focus



marker, and not as a classifier. Scholars have still not reached a consensus as for the status of such particle.<sup>99</sup>

- 424) Hai cái con chó y đâu?  
two cái CL dog that where  
'Where are those very two dogs?'

With respect to modifiers, we should note that there is another way to distinguish between classifiers and measure phrases. According to Nguyen (2004, 2008), a measure phrase can be modified, a classifier cannot. The following example contains a modified measure phrase (2008:17):

- 425) Một bình Nhật trà Tàu.  
one pot Japanese tea Chinese  
'One Japanese pot of Chinese tea'

In the example above, the cluster *bình + Nhật* forms the modified measure phrase.

If a numeral is present, it will precede the CL+N phrase and the CLP cannot be topicalized at the left periphery. This means that the numeral cannot be left alone. This is, however, possible with bare nouns. We will see this in more detail in the section about bare nouns in Vietnamese.

Modifiers have to occur to the right of the noun, and cannot separate the classifier from the noun. A measure noun, on the other hand, may be modified forming a measure phrase (without classifier).

As for the category number, Vietnamese does not have an overt way to mark the contrast singular vs. plural. As a matter of fact, a bare noun may yield both singular and plural readings, both definite and indefinite interpretations. If a classifier is present, then the entity is always singular because of the singularizing function of classifiers. We will see these facts in more detail below, in the section about Vietnamese bare nouns.

There are, however, ways to overtly mark plurality, for example via plural expressions such as *những* 'some of' or *các* 'the.PL' (2008:19).

---

<sup>99</sup> Cái has been previously analyzed as a general classifier, an identifier or an indexical *cái*; Nguyen C. (2012) describes it as a partitivity marker.

- 426) Những cuốn sách hiếm.  
 PL CL book rare  
 ‘(Some of the) rare books.’
- 427) Các con ngựa đen.  
 PL CL horse black  
 ‘The black horses.’

There is a further grammatical category, namely the determiner, which we have found in each language taken in comparison in this chapter, with the exception of Mandarin Chinese.

Vietnamese is usually referred to as a language without determiners. To the contrary, Nguyen (2004, 2008) assumes that determiners in Vietnamese do exist.

The numeral *một* would act as an indefinite determiner. According to Nguyen (2004, 2008), a classifier in Vietnamese may yield both a definite and an indefinite interpretation, but if the [CL-N] phrase is introduced by the numeral *một* ‘one’, it will always be singular and indefinite. As remarked by Nguyen, *một* only occurs with a noun if it is the first mention of that noun. It is consistent with the association indefiniteness and new information. If it is not the first mention, then a classifier is used for introducing the noun. Such facts lead Nguyen to claim that *một* is a singular indefinite determiner (2008:33).

- 428) Bà tôi cho tôi một con mèo.  
 Grandmother I give I one/a CL cat  
 ‘My grandma gave me a cat.’
- 429) Đó là một con chó.  
 That be one/a CL dog  
 ‘It is a dog.’

According to Nguyen, there are two further determiners in the nominal system of Vietnamese, namely a definite and indefinite one, which we already met above. They are, indeed, the plural markers *những* and *các*. The latter is a plural definite determiner, whereas *những* is the indefinite plural article. The determiner *những*, according to Nguyen (2008:39), refers only to a part of the whole set, whereas *các* denotes all entities of the set. It does not necessarily entail (non)specicity. The following sentences exemplify the occurrence of *những* (2008:40f.):

- 430) Bà tôi nuôi nhiều mèo. Những con mèo đen...  
 Grandma I raise a lot cat. NHỮNG CL cat black  
 ‘My grandma has a lot of cats. Some of the black cats...’

- 431) Công ty này đang tuyển những công nhân có kinh nghiệm.  
 Company this PROG select NHỮNG worker have experience  
 ‘This company is looking for experienced workers.’

In the following sentence, *những* yields an existential reading:

- 432) Có những bài thơ đọc rồi nhớ mãi.  
 Have NHỮNG CL poem read already remember forever  
 ‘There are some poems that once read become unforgettable.’

As we already said above, *các* differs from *những* since it “refers to the totality of entities that can be uniquely identified by both the speaker and hearer, either from previous discourse or the shared knowledge” (Nguyen 2008:43). As the following sentences will show, whenever *các* introduces a noun, it will automatically yield a plural definite interpretation:<sup>100</sup>

- 433) Hãy bày các quân cờ ra.  
 IMP display piece chess out  
 ‘Set up the chess pieces.’
- 434) Cho tôi mượn các cuốn sách mới.  
 Give I borrow CÀC CL book new  
 ‘Lend me the new books.’
- 435) Các em (này) rất ngoan.  
 CÀC younger sibling this very well-behaved  
 ‘These children are very well-behaved.’

Crucially, Nguyen notes that “Vietnamese lacks an overt determiner to unambiguously mark [+ singular, + definite]” (2004:42).

---

<sup>100</sup> Nguyen (2008:44) remarks that whenever *các* co-occurs with kinship terms, “the resulting phrases refer inclusively to all the members of a class or group that are being addressed”. This is shown in the following example:

Các em (này) rất ngoan.  
 CÀC younger sibling this very well-behaved  
 ‘These children are very well-behaved.’

## 5.7.2 Vietnamese bare nouns

Bare nouns in Vietnamese may be found in both argument positions, i.e. external and internal arguments. This entails that we may find both preverbal and postverbal bare nouns, just as in the other languages analyzed so far in the present chapter.

In the preceding section, we followed Nguyen's claim that *những* and *các* are plural determiners. If we assume this, then they are not (prenominal) plural markers and, therefore, there is no morphologically overt plural marker in Vietnamese. This is consistent with the claim that Vietnamese does not have overt number marking since it is an isolating language.

Moreover, the noun is not morphologically specified for singular or plural: it seems rather that it is nonspecified for number. Therefore, Vietnamese bare nouns may yield both singular and plural readings. The sentences in (436) and (437) will clarify this fact:

- 436) Tôi mua sách.  
I buy book  
'I bought (a/the) book(s).'
- 437) Tôi trông thấy hổ.  
I look see tiger  
'I saw (a/the) tiger(s).'

The sentences above may yield both a singular and plural interpretation. Interestingly, the bare nouns *sách* 'book' and *hổ* 'tiger' are ambiguous between a definite and an indefinite reading. It seems thus that the interpretation of Vietnamese bare nouns gives much space to the context.

The same ambiguity as for number and (in)definiteness interpretation is found in the preverbal position. The following example from Kirby (2006) will show this:

- 438) Bò đang ăn lúa kia.  
cow PROG eat paddy over-there  
'(A/the) cow(s) is/are eating (your/the) paddy (over-there)!'

As noted in Kirby, if *bò* 'cow' in (438) were introduced by the [+animate] classifier *con*, the CIP would receive a singular reading. The ambiguity between a definite and an indefinite reading would remain, just as in the case of bare nouns.

As we saw above with regards to number, classifiers have the function of ‘individuating’ and making the noun countable. They are therefore singularizers, and a noun introduced by a classifier will always yield a singular interpretation. As a consequence, if a classifier had introduced *sách* in (436) or *hổ* in (437), the only possible reading with respect to number would be the singular. On the other hand, the (in)definiteness status would, however, remain ambiguous, needing contextual support. Take a look at the following example:

- 439) Tôi muốn mua cuốn sách.  
I want buy CL book  
‘I want to buy a/the book.’

The possible readings are definite or indefinite, but always singular. However, according to Kirby (2006), whenever the postverbal noun is introduced by a classifier, it will yield a specific interpretation.

If we want to yield a plural reading, we need to leave the classifier out, as in the following examples (2008:18):

- 440) Sách rất nhiều.  
Book very a lot  
‘There are a lot of books.’
- 441) Chó cắn nhau.  
Dog bite each other  
‘Dogs fight against each other.’

The bare nouns *sách* in (440) and *chó* in (441), both syntactic subjects in the sentence-initial position, yield plural readings. The former occurs in an existential sentence, whereas the latter is the subject of a transitive predicate in a reciprocal construction.

By contrast, Trinh (2011) claims that Vietnamese bare nouns cannot be definite. They would yield a generic interpretation rather than a definite one. The generic interpretation of the Vietnamese bare noun is also taken into account in Kirby (2006). The example below from Trinh (2011) shows the preference for a generic reading:

- 442) Cho thích ăn thịt.  
Dog like eat meat  
‘Dogs / \*The dog(s) like(s) to eat meat.’

Trinh explains this by assuming a preference principle, for which whenever a kind reading may be yielded, then it is preferred to the definite reading, under the condition that there is no type mismatch. This is exactly what Chierchia (1998) proposes, namely the preference of the kind operator over the definite article.

As we anticipated in the preceding section with respect to the possibility of topicalization in Vietnamese, nouns in a CIP cannot be topicalized at the left periphery if they are introduced by a numeral. Topicalization of the noun would result in a numeral without another constituent. Such possibility is excluded by the Vietnamese grammar. It is, however, possible to move the noun, leaving its classifier with the numeral in its usual position. A bare noun would, thus, result in the topic position. This is shown in the example below:

- 443) Sách tôi có hai cuốn.  
book I have two CL  
'As for books, I have two.'

In sum, we find in Vietnamese the following possibilities of interpretation for bare nouns. As for number, it can yield both singular and plural readings. On the other hand, with respect to the category (in)definiteness, bare nouns may be definite or indefinite. It seems, thus, that much work for the interpretation is left to pragmatics and to the discourse context. Furthermore, bare nouns in Vietnamese also allow a generic reading.

As we saw in the preceding section, both CLPs and DPs may be found in Vietnamese. With respect to CLPs, they always yield a singular interpretation. On the other hand, they are not semantically specified for (in)definiteness. This implies that both definite and indefinite interpretations are possible for CLPs. Finally, we saw that there are three lexical items that can be considered as determiners: the indefinite singular *một*, the indefinite plural *những* and the definite plural *các*. The first two, i.e. *một* and *những*, always yield an indefinite interpretation, singular and plural, respectively. The crucial difference is that *một* may be used for the first mention of the referent, only. Last, *các* always yields plural definite readings since there is no singular counterpart of the definite determiner.

## 5.8 Gbe languages

In this section we are going to look at the nominal system(s) of Gbe languages, a branch of the Kwa family. Kwa languages are spoken in western Africa. The linguistic situation of West Africa is a bit too complex to be dealt with in this section. Nonetheless, I would like to simply shed some light on the Kwa cluster, in order to better define the object of study of the present section.

According to Capo (1991), based on Greenberg's (1966) classification of African languages, the Kwa family belongs to the Niger-Congo branch of Niger-Kordofanian. According to Aboh (2010), Kwa languages, formerly Eastern Kwa, go back to a Proto-Kwa and split into the following groups: Potou-Tano (consisting of the two clusters Potou and Tano, which further split into a number of other subgroups), Na-Togo, Ka-Togo, and Gbe. Each of these language groups further splits into several other subgroups. We will focus on the Gbe group.

According to Capo (1991) and Aboh (2004), the Gbe cluster consists of five subgroups: Vhe (or Ewe), Gen, Ahá, Fon, and Phla-Pherá. In Aboh's words, "it comprises all languages or dialects that refer to 'language' by using the lexeme *gbe*" (2004:24; my emphasis). According to Capo (1991) and Aboh (2004), these languages are spoken in the southern part of the most eastern region (Volta) of Ghana, in the southern part of Togo and Benin, and in different areas of the Nigerian states of Ogun and Lagos. The languages belonging to the Gbe cluster seem to be quite homogeneous with respect to each other as for their phonology, morphology and syntax. These facts suggest a common origin for these languages.

Before concluding this introduction on the Gbe cluster, we need to say that these languages are typologically different with respect to the languages analyzed so far in the present study. Gbe languages (or at least some of them, such as Gungbe and Fongbe) have, indeed, case markers, which we do not find in the other languages taken into account so far. At least, they have case markers for the genitive.<sup>101</sup>

---

<sup>101</sup> For instance, Fongbe has two case markers, *sín* and *tòn*. The former marks the objective and is translatable as 'of'; the latter marks the subject for the genitive case and may be rendered as 's'. A further distinction between both markers is in their distribution: *sín* is prenominal, whereas *tòn* is postnominal (Lefebvre and Brousseau 2002). Also, Aboh (2004:132f.) speaks about the genitive case marker of Gungbe, which seems to be syncretic with the objective form in Fongbe, i.e. *sín*. Gunge *sín* is prenominal as in Fongbe. The following examples from Fongbe will clarify the use of such case markers: *didè Kòkú tòn le*, sketch Koku Case PL, 'Koku's sketches' vs. *Koku sín dide le*, Koku Case sketch PL, 'sketches of Koku' (Brousseau and Lumsden 1992:7).

On the other hand, Gbe are typologically “close” to Chinese, which was dealt with in the preceding section, since both are tone languages, as a crucial difference from the creole languages analyzed so far. Gbe languages have two tones, namely high and non-high.<sup>102</sup> Finally, according to Aboh (2004), Gbe languages are SVO languages, with the exception of imperfective and related constructions.<sup>103</sup>

We will see Gungbe and Fongbe in more detail, both belonging to the Gbe subgroup Fon.

### 5.8.1 The nominal system of Gbe languages

In these languages, there is no morphological gender distinction. Moreover, nominal inflectional morphology is quite poor. For instance, there is a way to mark the noun for number (i.e. plurality), but nonetheless the noun is usually bare. We will see in more detail below under which conditions the plural marker is found. Furthermore, whenever a numeral accompanies the noun, this latter is bare. This is the same (tendential) pattern we found in the creoles analyzed in the present chapter. Furthermore, determiners, demonstratives, and modifiers do not agree in number (nor in gender) with the noun they modify. They do not inflect, indeed, either for number or for gender.

According to Aboh (2004, 2010), Gbe languages present head-complement structures (2004:3), although they superficially seem to be head-final, at least in certain contexts. Determiners and demonstratives, in fact, occur at the right of the noun. This yields the superficial head-final structure. The same is true for postnominal phrases.<sup>104</sup> The fundamental word order in Gungbe as for the nominal system is shown in Figure 4 and is exemplified by the Gungbe example in (444):

Figure 4

Noun – modifier – numeral - demonstrative – determiner - number

---

<sup>102</sup> According to Aboh (2004), tone languages need both pitch phonemes (tonemes) and segmental phonemes in the composition of morphemes (2004:27). In Gbe languages, tones can be of two types: lexical and syntactic. A tone belongs to the former type when it has an influence on the lexical level. On the other hand, syntactic tones also play a grammatical role. They can, in fact, express force, tense, mood, and aspect (2004:27-30).

<sup>103</sup> Aboh (2004) describes that the internal argument of the verb in the imperfective construction in Gungbe moves to the preverbal position, which yields the superficial order SOV (Aboh 2004:194-198).

<sup>104</sup> Aboh (2004:115-123) uses the term ‘postnominal phrases’ for items, which look like postpositions. His purpose is to distinguish them from real postpositions. The latter are found in languages such as Hindi and Japanese, among others. In Gungbe, on the other hand, postnominal morphemes would act as “(nominalizer) heads and force the DP they select to appear in case or governed positions” (Aboh 2004:120).



- 444) Àvún wéwé àwè èhè ló lé.  
 Dog white two DEM DET PL  
 ‘These two white dogs’.

As for determiners, as we said above, they always follow the noun. The following examples from Gungbe and Fongbe will clarify these facts (Aboh 2004:31):

- 445) xwé ló (Gungbe)  
 xwé ó (Fongbe)  
 house D  
 ‘the house’

Such postnominal determiners have the main function of marking a noun as specific. If the noun is nonspecific, then we have a bare noun. As we will see below, bare nouns occur quite freely in Gbe languages and may yield both definite and indefinite interpretations, mostly depending on the context. As for (non)specificity, they are always nonspecific since specificity relies on overt determiners.

Consequently, Aboh (2004) refers to Gbe determiners as specificity markers. In Gungbe there are two different specificity markers. The difference lies in that the one marks the noun as specific definite, and the other as specific indefinite. The following examples from Gungbe (Aboh 2004) will clarify these facts. In (446) we have the specificity marker *ló*, yielding definiteness. On the other hand, in (447) we have *dé*, the indefinite specificity marker.

- 446) Kokú mon távò cé bó do émi ná xo távò ló.  
 Kolu see-Perf table 1sg-Poss and say-Perf 3sg-Log Fut buy table Spec<sub>[+def]</sub>  
 ‘Koku saw my table and then said he would buy that specific table.’
- 447) Kokú mon távò cé bó do émi ná xo távò dé.  
 Kolu see-Perf table 1sg-Poss and say-Perf 3sg-Log Fut buy table Spec<sub>[-def]</sub>  
 ‘Koku saw my table and then said he would buy a specific table.’

As in many other languages, these two specific markers are in complementary distribution. Thus, their co-occurrence is excluded from Gbe grammar. So, what we have seen so far for many languages, namely that their noun phrases, bare or not, may be ambiguous between a specific and nonspecific interpretation, does not hold for Gbe languages, whose specific status is always expressed via postnominal markers, further specified for (in)definiteness. Whenever

such markers do not follow the noun, then the noun will be nonspecific. Thus, ambiguity may arise in Gbe languages for the (in)definite status of bare nouns, only.<sup>105</sup>

The determiner system of Fongbe is quite similar to the Gungbe case. Following Lefebvre and Brousseau (1992), *ó* is the definite determiner and seems to imply specificity, just as in the case of Gungbe. It is not specified for gender, therefore it may occur with both male and female, animates and inanimates. In contexts where a nasal consonant occurs, we find an allomorph of the definite determiner, i.e. *ón*. The use of *ó* and its allomorph are shown in (448) and (449), respectively.

- 448) *Ví ó.*  
 Child DEF  
 ‘The child.’
- 449) *Àvùn ón.*  
 dog DEF  
 ‘The dog.’

Both determined nouns above yield, thus, a specific interpretation. According to Lefebvre and Brousseau, it is not possible for the definite determiner to introduce nouns that yield generic or mass interpretations. They assume the definite determiner to be always anaphoric. In case of a mass noun with a postponed determiner, it will yield a taxonomic reading or in the author’s words (2002:38), a specific distributive interpretation of the kind ‘type of x’. This is shown in sentence (450). Furthermore, between the noun and its determiner, a relative clause may occur (451).

- 450) *Mòlikún ó.*  
 rice DEF  
 ‘The type of rice’.
- 451) *Súnû dé-é yì ó.*  
 man OP-RES leave DEF  
 ‘The man who left’.<sup>106</sup>

---

<sup>105</sup> Following Lefebvre and Brousseau (2002), the definite determiner *ó* in Gungbe has also another function: it may also appear in simple clauses. In such cases, it occurs at the end of the sentence and “presupposes shared background information” (2002:481). It yields scopal ambiguities, basically deriving three types of shared information. It could be something already known in the discourse context, it could refer to the whole event denoted by the clause, or to only a subpart of the event. In the latter two cases, it is referred to in the literature as an ‘event determiner’ and may trigger a subject-oriented and an object-oriented interpretation (2002:481-502).

<sup>106</sup> In (399), OP-RES stays for the cluster operator + resumptive pronoun: the operator pied-pipes the resumptive pronoun (Lefebvre and Brousseau 2002:38). Furthermore, Lefebvre and Brousseau specify

The second determiner we find in Fongbe is the indefinite *dé*, which Lefebvre and Brousseau assume to be presumably a reduced form of the numeral *óde* ‘one’ (451). This indefinite determiner occurs only with specific nouns, whereas nonspecific indefinite noun phrases are expressed via bare nouns. The indefinite determiner may also be postponed to a mass noun. In this case, the mass noun yields a taxonomic interpretation, as in the case of mass nouns with the definite determiner (see above).<sup>107</sup>

As for plurality in Gbe languages, it can be overtly expressed by the number marker. This marker is also postnominal, but whenever a specificity marker is present, then the number marker has to occur to the right of the specificity marker. Thus, the number marker occupies the rightmost position. Interestingly, according to Aboh (2004:81), the number marker in Gbe languages might have derived from the third plural person pronoun. This would be similar to the cases of Santome and Papiamentu, as we already saw, and, following Aboh, these facts would be quite transparent in Gengbe and Ewegbe. In these languages, the number marker and the weak pronominal form for the third plural person pronoun are syncretic (*wó*).

In the sentence (452) below, we find the Gungbe plural marker *lé* combined with the specificity marker(s). Important to note is that the plural marker bears some degree of specificity, or at least it is found either in combination with the specificity markers, or if no specificity marker occurs, it yields a specific interpretation (76-84).

- 452) Koku xo            tavo            lo/de l lé.  
 Koku buy\_Perf table Spec<sub>[+def]</sub>/Spec<sub>[-def]</sub> Num  
 ‘Koku bought the/some specific tables.’

On the other hand, the Fongbe plural marker *lé* is not syncretic with the third plural person pronoun, which has the form *yé*. Furthermore, the Fongbe plural marker is inherently specified for definiteness, as we can understand from the example in (453):

- 453) Àsón lé.  
 crab PL  
 ‘The crabs’.

---

that *dé* is a nominal operator since it can pied-pipe nouns, but not postnominal phrases of the type *távò jí*, ‘table on’. The operator *dé* is found in both relative and factive clauses. See Lefebvre and Brousseau (2002:118f.) for further discussion on this topic.

<sup>107</sup> *Dé* is also used for the compound form *déó* ‘the other’ and *dévó* ‘another’ (Lefebvre and Brousseau 2002:40).

The overt pluralized noun in (453) yields a definite interpretation. An indefinite plural reading such as ‘some crabs’ is excluded from Lefebvre and Brousseau (2002:39). Following the authors, this would distinguish Fongbe from, for instance, Ewe. Ewe plural marker *wó*, indeed, allows both definite and indefinite interpretations. Unlike Ewe, indefinite plurals in Fongbe are expressed via bare nouns. On the other hand, if the indefinite plural is specific, Fongbe uses the cluster indefinite determiner *dé* + the plural marker *lé*, roughly translatable as ‘some, a few’.

Furthermore, in Fongbe just as in Gungbe, as we saw above, the definite determiner may co-occur with the plural marker, which will be the most rightward element of the DP:

- 454) Àsón ó lé.  
 Dog DEF PL  
 ‘The crabs.’

Another interesting phenomenon of the Gbe grammar is the topicalization. According to Aboh (2004), along with focalization and *wh*-constructions, topicalization is a phenomenon regarding the left periphery of the Gbe sentence (51). Rizzi (1997) defines a topic as “a preposed element characteristically set off from the rest of the clause by a ‘comma intonation’ and normally expressing old information somehow available and salient in previous discourse (Rizzi 1997:285, in Aboh 2004:51). According to Aboh, in Gbe there is left dislocation of the topicalized noun, which, in some Gbe languages, is followed by a topic marker and leaves a resumptive pronoun in the IP-internal position. The topic precedes the focus, according to Aboh (2004:51). The following sentences illustrate these facts in Gungbe (455) and Fongbe (456), respectively (*ibid.*):

- 455) ...dò dàn ló yà Kòfi wé hù-ì  
 that snake DET TOP Kofi FOC kill-Perf-it  
 ‘...that the snake, Kofi killed it.’

- 456) ...dò dàn ó, Kòfi wé hù-ì  
 that snake DET, Kofi FOC kill-Perf-it

The crucial difference between the two sentences above is that the Gungbe sentence in (455) marks the topic *dàn* ‘snake’ via the topic marker *yà*. On the other hand, in the Fongbe sentence (456), the topic is not highlighted by a topic marker.

With respect to means for expressing deixis and specificity, it is worth noting that Fongbe shows an interesting alternation in the interpretation of its demonstratives. There seems to be a bipartite system with three possibilities of interpretations. There are two demonstrative forms, namely *ló* and *né*, each of them having an emphatic form *éló* and *éné*, respectively. According to the three different patterns of interpretation described in Lefebvre and Brousseau (2002), (*é*)*ló* can be interpreted as proximal in the first system and as both proximal and distal in the other two. On the other hand, (*é*)*né* is distal in the first and second pattern, but it may be both proximal and distal in the third. As in the case of determiners, demonstratives do not undergo any morphological (nor phonological) change, as for gender and number. As we saw above for Gungbe in Figure 3, demonstratives in Fongbe are also postnominal and occupy a middle position between modifiers and determiners. Finally, the demonstrative determiner cannot co-occur with the indefinite determiner. They are in complementary distribution. These facts are shown in the example below (from Lefebvre and Brousseau 2002:42):

- 457) Àlòké cé      éló/éné ó      lé.  
ring POSS DEM/DEM DEF PL  
‘These/those//these/those rings of mine.’ (specific)
- 458) Ùn mò      éló/éné.  
1sg see DEM/DEM  
‘I saw this/that//this/that one.’

In (457) we see the possible combinations of the demonstratives with a possessive, with the definite determiner and the number marker. On the other hand, in (458) we can see the pronominal use of the demonstratives.

As for modification, we saw in Figure 4 above that the modifier is postnominal and occupies a middle position between the noun to its left and the demonstrative to its right.<sup>108</sup> These facts are exemplified in the Gungbe sentence in (459) – (Aboh 2010):

- 459) Àsé yù éhè.  
cat black this  
‘This black cat’.

---

<sup>108</sup> When the demonstrative directly follows the noun and precedes the modifier, then we are not dealing with a modified demonstrative phrase, but rather with a predicative construction: *àvún éhè kló*, dog DEM big, ‘this dog is big’.

More on adjectives, there seem to be in Gungbe (and in general, in Kwa languages- Aboh 2010) two classes of adjectives, just as we saw in the case of Kriyol. The difference would lie in their function, namely attributive vs. predicative. To the first class there belong attributive adjectives. They seem to be quite few and “often denote color, size, and shape” (2010:14). In (459) above, we saw an example of such adjectives in its attributive function, i.e. *yù* ‘black’. On the other hand, items with predicative function are referred to as “adjectival verb constructions” by Aboh (2010:14f.). They differ from real adjectives in a number of respects: first, for their position with respect to the noun; and second, adjectival verbs may combine with TMA markers, as any lexical verb, and again similarly to verbs, they can undergo (focalizing or relativizing) fronting (with doubling). Furthermore, attributive adjectives need a copula in order to be used for predications, whereas verbal adjectives do not. This is the very same situation found in Kriyol. Furthermore, adjectival verbs show a different morphological behavior whenever they are used attributively. More specifically, they reduplicate and occupy the same position as real adjectives.

We find a similar situation in Fongbe, according to Lefebvre and Brousseau (2002:348): there seems, indeed, to be very few real adjectives in Fongbe. The authors claim that in contexts where English would use an adjective, Fongbe uses stative verbs: *bí* ‘to be alert’, *kló* ‘to be big’, *wíni* ‘to be slim’, etc.

- 460) Kòkú bí / dá / kló.  
 Koku be.alert / be.bad / be.big  
 ‘Koku is alert/bad/big.’

With regards to predicative construction, as remarked by Lefebvre and Brousseau, nouns and real adjectives in predicative positions must be introduced by a copula. There are two copulas which can relate the subject to the predicative item, namely *nyí* ‘to be’ and *dò* ‘to be at’.

- 461) Kòkú nyí mé sí  
 Koku be teacher  
 ‘Koku is a teacher.’
- 462) Koku dò dàgbè.  
 Koku be.at good  
 ‘Koku is good.’

As in Gungbe, real adjectives<sup>109</sup> in Fongbe, unlike stative verbs, can modify nouns. Via partial reduplication, adjectives may be derived from verbs.<sup>110</sup> On the other hand, real adjectives do not undergo such a process (Lefebvre and Brousseau 2002:350).<sup>111</sup>

Finally, as for numerals, they are also postverbal. As we saw in Figure 4 above, whenever a numeral co-occurs with a modifier, the numeral will follow it, as shown in (463).

- 463) Àwó wéwé àwè. (Gungbe)  
Cloth white two  
'Two white clothes.'

Whenever a numeral occurs to modify a noun in Gbe sentences, the noun will be bare. Only if specificity is implied, the noun will be marked for number, as we saw in the sentences in (452-454) above.

According to Aboh (2010:23), it holds generally for Gbe languages that the number marker is never allowed in DP-initial position: \**wo ete eya*, PL yam DEM (Gengbe), or \**ɛ'tèví éhè*, PL yam DEM (Gungbe). Number-marked demonstratives are also excluded in Gbe languages (Aboh 2010:23).<sup>112</sup>

### 5.8.2 Bare nouns in Gungbe and Fongbe

According to Aboh (2010), Kwa languages in general can use bare nouns in every context, yielding a wide range of interpretations, although there are differences among the languages belonging to this family. As in the previous section, we will look in more detail at the case of Gbe languages, with particular emphasis on Gungbe and Fongbe.

Following Aboh, bare nouns in Gungbe occur quite freely and may occupy both the subject and object positions. As we said above, bare nouns are always nonspecific in Gbe languages since specificity is realized via overt specificity markers, further specified for

---

<sup>109</sup> Along with real adjectives, Fongbe has also a small group of 'adjectivized nouns', which are mostly derived by means of the suffix *-nò*: *àkwé'nò*, 'rich', *àdògónò*, 'portbellied' etc. (Lefebvre and Brousseau 2002:354).

<sup>110</sup> A further distinction between real adjectives and nouns used as adjectives, on the one hand, and stative verbs, on the other, is that the latter but not the former can undergo partial reduplication: *Bí*, 'be alert' < *bíbí* 'agile/alert/dynamic'. This is a process for deriving adjectives from verbs (Lefebvre and Brousseau 2002:349).

<sup>111</sup> Some real adjectives can undergo full reduplication, and the result is that of intensification (Lefebvre and Brousseau 2002:351).

<sup>112</sup> In some Gbe languages, the demonstrative may be separated from the number marker also by a determiner, more precisely by a specificity marker (Aboh 2010:23f.).

(in)definiteness. Ambiguity may rise just in case of the (in)definite status of bare nouns in these languages. Following Aboh, the (in)definite interpretation of bare nouns depends primarily on the context.

As for the subject position, a bare noun in Gungbe may yield both definite and indefinite interpretations, mostly depending on the context.

The same is true for bare nouns in the object position. Both definite and indefinite readings are possible. It seems, thus, that the correct interpretation depends on the context. Nonetheless, I would not exclude that the general tendency for these languages is to have a definite interpretation in the subject, and an indefinite one in the object position.

In (464) and (465) below, we have bare nouns in the subject position in Gungbe and Ewegbe, respectively. The bare subject in the Gungbe sentence (464) is interpreted as indefinite since it is new information, answering the question “what happened?”. On the other hand, in the Ewegbe sentence in (465) a bare noun occurs in the subject position and yields a definite interpretation, easily explainable by the fact that it is a unique entity:

464) Àsé jε` càzù mé. (Gungbe)

Cat fall pot in

‘A cat fell in a pot!’

465) Ddo vu sesie egbea akpa. (Ewegbe)

Sun open hard today too.much

‘The sun was too hot today.’

As for the object position, the Gungbe sentences below in (466) show the tendency for the bare object to be interpreted as indefinite. On the other hand, the sentence in (467), also from Gungbe, will show that, depending on the context, it is still possible for a bare object to yield a definite interpretation (following Aboh, the sentence in (467) was uttered in a context where only one doctor was present):

466) Kpón! Mús jró ná wlé kòkló.

Look, Mus want PREP catch fowl

‘Look! Mus is trying to catch a fowl!’

467) Bé à món dòtó tò dôn?

Q 2sg see doctor at there

‘Did you see the doctor here?’



Moreover, a bare noun in the internal argument position is not limited to a singular indefinite interpretation. According to Aboh, depending on the context, a bare object may also be interpreted as plural indefinite, as shown in the Gungbe sentence in (468), where the bare object *àsé* ‘cat’ may yield both a singular indefinite and a plural indefinite interpretation.

- 468) Ùn jéyì àxìmé bò ná yì xó àsé.  
 1sg going market COORD FUT go buy cat  
 ‘I am going to the market to buy a cat/cats.’

So far, we have seen that neither the subject nor the object position need an overt specification of their (in)definite status. According to Aboh, the same holds for focused, questioned and relativized nouns. In other words, bare nouns freely occur as focus, in questions and followed by a relative clause, and may yield both definite and indefinite interpretations. The following Gungbe examples from Aboh (2010) will show these facts:

- 469) Àsé wé Kòjò zé hwèví blébù ná  
 cat FOC Kojo take fish whole PREP  
 ‘Kojo gave a whole fish to a/the cat!’
- 470) Àsé tɛ’wé Kòjò zé hwèví blébù ná?  
 Cat Q FOC Kojo take fish whole PREP  
 ‘Which cat did Kojo give a whole fish to?’
- 471) Àsé dǝ Kòjò zé hwèví blébù ná.  
 cat REL Kojo take fish whole PREP  
 ‘The cat which/that Kojo gave a whole fish to?’

Furthermore, bare nouns in Gungbe may also yield generic readings, both singular and plural:

- 472) Ùn nyín wán ná àsé.  
 1sg COP sentiment PREP cat  
 ‘I love cats.’

Roughly speaking, we may say that the correct interpretation of bare nouns in Gbe languages often depends on the context, following Aboh (2010:11-14).

As for bare nouns in Fongbe, they may be found in generic statements, yielding a generic reading. This is shown in the sentence in (473) below:

- 473) Wòxúnú nyó nú lánmèyén.  
 Bread be.good for health  
 ‘Bread is good for one’s health.’

Bare nouns in Fongbe may also occur in episodic statements; however, they always yield a nonspecific interpretation. This is similar to the case of Gungbe. Both in Fongbe and Gungbe, in fact, specificity is always expressed via the definite determiner or via a demonstrative.

Crucial in Fongbe is the distinction mass vs. count nouns. Whenever the denotation of a noun is mass, the noun will be bare, as the example in (474) will show.

- 474) Koku du bledi.  
 Koku eat bread  
 ‘Koku ate bread.’

On the other hand, the case of count nouns is slightly different. Once again, the crucial factor here is the (non)specificity. Whenever nonspecific, both singular and plural count nouns are expressed via bare nouns. In case the noun is semantically plural, it will be morphologically marked for plurality only if it is specific. In this case, the plural marker *lé* will follow the noun. On the other hand, indefinite nonspecific plurals are expressed via bare nouns, as the sentence in (475) will clarify.

- 475) Ûn xò àsón.  
 1sg buy crab  
 ‘I bought some crab.’

Concluding, Gungbe and Fongbe seem to behave quite similarly as for the distribution and the interpretation of bare nouns. They may occur in both generic and episodic statements and occupy both external and internal argument positions. As for their interpretation, bare nouns are always nonspecific; as for their (in)definite status, the context is crucial. Finally, they may yield both singular and plural readings.

## 5.9 Summary and comparison: BNs among non-creoles

In their distributional properties, Mandarin Chinese, Vietnamese and Gbe (Gungbe and Fongbe) are very similar. In all these languages, in fact, bare nouns are allowed in both

the subject and object positions. The same similarity should be noted as for number interpretation. BNs unspecified for number in these languages may yield both singular and plural interpretations. The facts dealt with so far are summarized in Tables 8 and 9 below.

As for their semantic interpretation with respect to (in)definiteness, genericity, and specificity, the situation is highly homogeneous, as Table 10 below will illustrate. We are going to consider here bare nouns morphologically unspecified for number, only.

In Mandarin Chinese, bare nouns may be interpreted as definite, (non)specific and generic in both preverbal and postverbal positions, but the postverbal position may in addition yield an indefinite interpretation, which the preverbal position does not allow. Thus, the difference between the two syntactic positions lies in their (in)definite interpretation.

As for Vietnamese, there seems to be no difference at all between the preverbal and the syntactic positions. Both may yield (in)definite, (non)specific and generic readings.

In the Gbe languages taken into account in the present study, namely Gungbe and Fongbe, the situation is quite homogeneous. The specific interpretation is the only one that makes a difference here. In more detail, bare nouns in Gungbe may be (in)definite, (non)specific, and generic in both the preverbal and postverbal position. In the object position, however, information structure intervenes in making the object position more plausible for an indefinite reading, i.e. locus for new information. On the other hand, in Fongbe both preverbal and postverbal positions may yield (in)definite, generic and nonspecific readings. It is important to note here that a specific reading does not seem to be available for count nouns. They will be, in fact, overtly marked for specificity.

In conclusion, there seems to be a high degree of homogeneity as for both the syntactic and semantic behaviour. From a crosslinguistic perspective, the only difference seems to be the absence of a specific reading for bare nouns in Fongbe in both the subject and object position. Recall that in this language specificity relies on overt markers such as demonstratives. The other differences are, on the other hand, language-internal variations. Mandarin Chinese does not allow the indefinite reading in the preverbal position. The last difference noted in the present section is the preference of the object bare noun in Gungbe for an indefinite reading. This is consonant with the usual information packaging of languages in general.

Table 8. *Syntactic distribution of BNs morphologically unspecified for number*

	Subject	Object
Mandarin Chinese	Yes	Yes
Vietnamese	Yes	Yes
Gungbe, Fongbe	Yes	Yes

Table 9. *Interpretation of BNs –[number] as for semantic number*

	Singular	Plural
Mandarin Chinese	Yes	Yes
Vietnamese	Yes	Yes
Gungbe, Fongbe	Yes	Yes

Table 10. *Possibilities of interpretation for BNs –[number]*

	Preverbal position	Postverbal position
Mandarin Chinese	Definite; generic; (non)specific	(In)definite; generic; (non)specific
Vietnamese	(In)definite; generic; (non)specific	(In)definite; generic; (non)specific
Gungbe	(In)definite; generic; (non)specific	(In)definite <sup>113</sup> ; generic; (non)specific
Fongbe	(In)definite; generic; nonspecific	(In)definite; generic; nonspecific

<sup>113</sup> The indefinite reading is preferable for the object position.

## Chapter 6

### Semantic number: Theoretical implications on bare nouns

#### 6.1 Depréz's Plural Parameter

We already introduced in Chapter 3 that among the studies on bare nominals focusing on both their semantic denotation and syntactic structure (Longobardi 1994, Chierchia 1998, among others), we also find the model built by Depréz (2005, 2006, 2007). We briefly discussed this model in Chapter 3, and now we will briefly review the general lines of Depréz's model as introduced in Chapter 3. Then, we will be able to discuss this model in more detail.

According to Depréz, with regard to the category 'number' and interpretation of bare nouns, there are two groups of languages, i.e. the +Plural languages, on the one hand, and the –Plural, on the other. The distinction between these two sets is based on a simple but crucial structural difference, namely whether their NumP obligatorily contains a semantic “counter” (2007:322) or not, thus giving rise to +PL and –PL languages, respectively. This semantic counter, as we will see later, introduces a number argument which calls for saturation (Depréz 2005).

Depréz (2007) claims that the basic denotation of nouns crosslinguistically is the kind denotation. Nouns are thus expressions of type <e>. Non-kind interpretations are derived by means of “the compositional combination of this basic denotation with a variety of operators introduced by the functional projections of a given language and ultimately by its morphology” (2007:320). This represents an important difference from Chierchia's (1998) Nominal Mapping Parameter. As we saw in Chapter 3, in Chierchia's model, indeed, noun denotation varies according to the ability of the noun in a specific language to be an argument, a predicate or both.

Furthermore, Depréz (2006) assumes number to provide a ‘criterion of individuation’ mapping a kind onto its realizations. For Depréz, nouns are not default mass. Count and mass are distinct realizations of kind.

Following Depréz, the syntactic structure of (bare) noun phrases would be essentially the same in both groups of languages. The Number Phrase, indeed, is always present. What crucially distinguishes the two sets of languages is the projection of the Number Phrase: NumP obligatorily projects in +PL languages, but not (necessarily) in –PL ones. Furthermore, the +PL group also requires the presence of a semantic counter in NumP. This semantic operator satisfies the number requirement, so that the noun can be interpreted as singular or plural without giving rise to any ambiguity. As a crucial difference, the semantic counter is not necessarily required in –PL languages. Consequently, Depréz correctly predicts that singular bare count nouns in the argument position are allowed in –PL languages only. Indeed, since +PL languages always require the projection of NumP, a singular bare noun would be uninterpretable and, thus, be excluded from the grammar.

As for +PL languages, Depréz (2007:321) claims that NumP must always be projected for both singularity and plurality. This implies that, as for countability, it is the semantic counter that makes it sure, at least in +PL languages. This latter is “a measure function which is the semantic translation of ‘countability’” (ibid.). In +PL languages, thus, we always have number specification for count nouns. The (bare) noun will be either singular or plural, positing that singular possibly means ‘= 1’ and plural ‘> 1’. This may sound trivial, but there are also ‘singular’ forms which are not real singulars. In fact, they are simply underspecified for number. This is the case of –PL languages, after Depréz’s notation; we will see these facts in more detail below.

To sum up, we have seen so far that +PL languages completely lack optionality as for NumP. It must be projected in +PL. As a consequence, the fundamental structure of noun phrases contains a NumP, which is always projected. This holds for bare nouns as well, both singular or (overt) plural or unmarked for number.

The situation as for –PL languages is different. They do not necessarily require the projection of NumP since their grammars do not always require the presence of a semantic operator. According to Depréz, in –PL languages we may have both the simple projection of NP and the projection of NumP. In case of projection of a simple NP, the result will be a bare singular. As we said in Chapter 3, the terminology ‘bare singular’ is not the most suitable. We used it to differentiate overt plural NPs. In fact, it would be more correct to speak of ‘bare nouns underspecified for number’. Following Depréz, bare nonplural nouns in –PL languages are instances of what Corbett (2000) calls ‘general number’. In other words, they are not real singular forms, but they are rather underspecified for number. As a consequence, they can

yield both singular and plural interpretations. For instance, verb morphology in GBC cannot help distinguishing whether the subject is singular or plural. In fact, we already noted in Chapter 2 that verbs in GBC do not present any inflectional morphology. Following Depréz, bare number-neutral NPs in –PL languages are instances of the ‘general number’.

As a difference, if NumP in –PL languages is projected, the semantic counter is ‘activated’ and needs to be saturated. According to Depréz, this can happen by means of overt plural morphology, numerals and an indefinite determiner. Plural morphology, numerals and the indefinite determiner would, thus, encode such a counter.

As we already noted, bare nouns in –PL languages, thus unmarked for number, are instances of the ‘general number’. This happens to be exactly the situation of Guinea-Bissau Creole and in creoles in general. Kihm (2007) also speaks of “unspecified for number” with respect to bare singulars in GBC. The underspecification with respect to number may be expressed as follows:

Figure 5      –[number]

Bare nouns with the feature represented in Figure 4 are able to yield both singular and plural interpretations.

From a typological perspective, creoles belong to the –PL group. On the other hand, languages such as English, French and Italian would be +PL. As predicted by Depréz’s parameter, such languages do not freely allow bare nouns in the argument position. Bare singulars are usually disallowed (they would, indeed, be uninterpretable), whereas bare plurals are allowed under certain restrictions. We will see below in more detail examples from both +PL and –PL languages, but before we do, we should look at the syntax of bare nominals and the possibilities Depréz proposes as for the derivation of interpretations of bare nouns in both sets of languages.

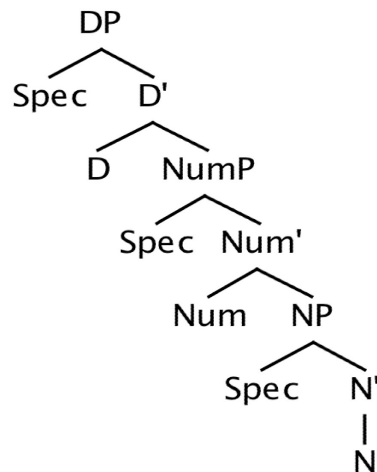
## 6.2 The syntax of bare nominals

As we said above, NumP is always present in the deep structure of every NP. The syntactic structure of bare nominals changes according to which denotation they have. More specifically, if they refer to kinds, they are supposed not to project the Number Phrase, i.e.

they are NPs. On the other hand, whenever they refer to realizations of kinds, they project NumP, which in turn contains a counter. In case it does not contain any counter, the nominal will be underspecified for Number, raising the ambiguity between singular and plural interpretation. This is the case of bare singulars in Kriyol.

As a consequence of Depréz's approach, the difference between a NP and a BNP would be that the latter has a null D, whereas the former has an overt D. In other words, both are full noun phrases with the same syntactic structure. According to Depréz (2005, 2007), partially based on Longobardi (1994), the syntactic structure of any NP (bare or not) would be as follows (Figure 6):

Figure 6



### 6.3 Derivation of BN interpretation within Depréz's account

We mentioned above that, crosslinguistically, kind is the basic denotation of nouns. In order to derive the other interpretations, the basic denotation combines with the operators that languages have at their disposal. Depréz assumes the semantic counter to be an operator acting over realizations of a kind. Hence, it turns kinds into objects. Following Depréz, the kind reading would be directly accessible for bare nouns in –PL languages only. On the other hand, +PL languages need a semantic operator to yield the kind reading. With this purpose in mind, Depréz proposes to use the GEN operator as discussed in Krifka et al. (1995). This does not seem, however, to be possible. The GEN operator has, in fact, a different role. More specifically, GEN has the function of providing the sentence with genericity. Following



Krifka et al. (1995), the GEN operator acts as a quantifier endowing the sentence with a generic interpretation. GEN can be defined as an operator generalizing over both objects and situations in characterizing sentences (Krifka et al. 1995:30-63). It is, therefore, not appropriate for yielding a kind-reading. This is a first problem within Depréz’s model. The derivation of kind-referring readings for bare nouns in +PL languages remains unsolved.

The situation is different when it comes to the nongeneric (object-referring) reading of nouns, namely when we have to derive objects from kinds. According to Depréz, in order to get an object from a kind in a –PL language, the Carlsonian realization rule applies to Num; this is shown in (476) below (Depréz 2007:324):

$$476) \quad \text{Num} = \lambda w \lambda x. \lambda K R_w(x, K)$$

As explained in Depréz (2007:324), “this realization rule is a relation from kind to objects that gives the set of objects  $x$  that instantiates a given kind  $K$  (where  $x$  is a member of the kind  $K$ ) in a given world  $w$ , independent of number”. Recall that this is the case of bare nouns in –PL languages, whenever the NumP fails to project. On the other hand, if the NumP projects and contains the required semantic counter, it will be this counter, together with the realization rule shown in (476), that turns the kind into a countable property. The same holds for +PL languages: Depréz assumes Krifka’s (1995) measure function OU (Object Unit) as a function returning a number whenever applied to a kind and an object. This is shown in (476) below. The number argument we get at this point must be saturated, “otherwise the resulting NumP fails to be interpretable” (2007:324). The saturation may happen via different items, e.g. a numeral, the indefinite determiner or plural morphology. In these cases, the number argument results “existentially quantified” (477).

$$477) \quad \text{Num} = \lambda n \lambda w \lambda x \lambda K [R_w(x, K) \wedge \text{OU}_w(K)(x) = n]$$

The formula in (477) has the following reading: “given a kind  $K$ , we obtain sets of objects  $x$  with cardinality  $n$  that are members (i.e. realizations) of the kind” (2007:324).

As for the existential reading of object-referring NPs, the two operations described so far, namely Carlson’s Realization rule and Krifka’s Object Unit, combine with operators such as the existential quantifier  $\exists$ . Following Depréz (2007), whenever  $\exists$  combines with  $R$  of (476), we will have the existential reading of bare nouns in –PL languages. On the other hand, in order to yield the existential reading of bare plurals in +PL languages, the combination of  $R$

and OU in (477) further combines with the existential reading. These operations are exemplified, following Depréz (2007), in the formulas below in (478) and (479), which represent the existential reading for bare nouns in –PL and for bare plurals in +PL languages, respectively:

$$478) \quad \exists x[R(x,DOG)]$$

$$479) \quad \exists x\exists n [R(x,DOG) \wedge OU(DOG)(x) = n > 1]$$

As for the generic reading, Depréz introduces the GEN operator, which combines with the above described operations R and  $R \wedge OU$  ((476) and (477), respectively). These facts are illustrated in (480) and (481) below for –PL and +PL languages:

$$480) \quad GENx [R(x,DOG)]$$

$$481) \quad GENx \exists x [R(x,DOG) \wedge OU(DOG)(x) = n > 1]$$

Recall that for both existential and generic readings, what is at stake is not just a bare NP, but a NumP. Within Depréz’s approach, in fact, NumP always projects. This explains the ban on bare singulars in +PL languages.

So far, the only reading which is not yielded is the kind-referring reading for bare plurals in +PL languages. Importantly, kind readings are directly accessible in –PL languages only. As a difference, since +PL languages always project number, they do not have direct access to kind readings, but to properties only. Depréz, thus, proposes that kind readings for +PL be yielded by means of operators. For this purpose, Romance and Germanic languages use an overt definite D. In English, the situation is different. The overt (plural) definite determiner does not yield kind readings. In Depréz’s proposal, this point remains unsolved and we are not able to yield kind readings for bare plurals in +PL languages.

## 6.4 Implications of the Plural Parameter:

### A comparison between +PL and -PL languages

We will examine in this section the predictions with respect to bare nouns of Depréz’s proposal. Following Depréz (2007), we will take into account a number of languages, both +PL and –PL. As for the +PL set, we are going to look at the case of English, French, and

Italian. On the other hand, among the –PL group Depréz examines a number of French-based creoles (Haitian, Mauritian, and Seychelles creoles) and a Portuguese-based creole (CVC). This last one represents a special case in Depréz’s classification. As we will see in more detail below, CVC happens to have a noun class [+human, +animate] which behaves as +PL. With respect to this peculiarity, we will assume that GBC is as special as CVC in this respect and will be, thus, treated on par with CVC.

#### 6.4.1 +PL languages: The case of English, French and Italian

The languages considered in this section will be English, French, and Italian. In these languages, bare singular count nouns are usually disallowed in argument positions. On the other hand, bare plurals are allowed under certain restrictions. Nonetheless, we have to point out that there are some exceptions, e.g. the names of professions seem to be freely allowed in the predicate position. This seems to be “the” exceptional case in all languages, both –PL and +PL (Depréz 2005:877). In the other cases of predicative use of bare nouns in +PL languages, namely nouns of the type <e,t>, these are to be treated as NumP within Depréz’s model. Just as with arguments, they call for number saturation. As a consequence, they occur either introduced by an overt determiner or with overt plural morphology. This is, however, the case of +PL languages.

As we already mentioned, bare singulars in the argument position are generally disallowed in +PL languages. On the other hand, +PL languages have bare plurals. They may yield both existential and generic readings. As for the kind reading, as we will see below, there are some important differences. The following sentences are examples of existential readings in English and Italian, respectively (adapted from Depréz 2007).

482) I heard dogs barking in the park yesterday.

483) I cani stanno giocando fuori.

(lit. ‘The dogs are playing outside.’)

In (484) and (485) below, we have examples of generic sentences. The English, French and Italian cases are treated as follows:

484) Dogs bark.

485) Leo corteggia sempre belle ragazze. (lit. ‘Leo always courts nice girls.’)

As for the ban on bare singular count nouns, it is predicted in Depréz's (2005, 2007) model: bare singulars in +PL languages are uninterpretable because NumP always projects in these languages and the counter needs to be saturated. This saturation may be achieved either by means of overt plural morphology or by an indefinite determiner. Whenever these conditions are not met, the measure function cannot result as saturated and, as a consequence, the bare noun is ruled out. In other words, in the case of singular bare nouns, the NumP contains a variable, namely its counter, which is unbound and, as a consequence, not interpretable. That is why singular noun predicates in English, French and Italian will always require an overt D in order to be licensed.

As we already said above, a kind reading is not directly accessible for bare (singular) nominals in +PL languages. In these languages, indeed, number morphology is always present in order to fulfill the measure saturation requirement. This seems to be the situation in Italian, as shown in examples (486) and (487) below; however, this seems to be possible in English (486).

486) Penguins are disappearing in Antarctica.

487) \*(Gli) elefanti sono estinti.

(lit. '(the) elephants are extinct')

To license *elefanti* as kind-referring in (487), we necessarily need the overt (plural) definite determiner. English, however, allows bare plurals as kind-referring.

Following Depréz, it seems that bare nominals in +PL languages need an operator, e.g. the definite determiner. This would be overt in French and Italian, and covert in English. Since this falls outside the purpose of the present thesis, we will not go through the arguments of Depréz. We simply limit ourselves at the description of the relevant facts on BN distribution (and restrictions) in +PL languages.<sup>114</sup>

---

<sup>114</sup> Depréz's (2005:877-880) solution is that of positing a null operator for deriving kind-referring readings for English bare plurals. This operator would be the one proposed in Krifka (1995), and its representation is as follows:

[<sub>NP</sub> bears]  $\sigma (\lambda w \lambda x \exists n (R/T_w (x \text{BEAR}) \& \text{OKU} (\text{BEAR})(x) = n))$   
 $= \iota y \forall w \forall x [\exists n (R/T_w (x \text{BEAR}) \& \text{OKU} (\text{BEAR})(x) = n) \text{ iff } R/T_w (x,y)].$

## 6.4.2 –PL languages: French-based and Portuguese-based creoles

In the present section, we will examine the case of –PL languages, and take into account the cases of French-based and Portuguese-based creoles. More specifically, we will first look at Haitian creole as described in Depréz (2006, 2007). Finally, we will deal with the case of two Portuguese-based creoles, namely Cape Verdean Creole and Guinea-Bissau Creole.

### 6.4.2.1 Haitian Creole

As described in Depréz (2005, 2007), Haitian Creole (henceforth HC) does not seem to have any restriction on the occurrence of bare nouns in the argument position. BNs can, in fact, show up in both the subject and object positions. As for number morphology, it is worth noting that there is no overt plural marker in HC: bare nouns, thus, always occur ‘underspecified’ as for number. There is a way to overtly mark plurality, i.e. *yo*. This plural marker, however, homophonous with the third plural pronoun, is inherently specified for definiteness. Depréz describes it as the plural definite determiner; its singular counterpart is *la*. These facts are clarified by the example in (488) below:

- 488) Jan achte liv la            vs. Jan achte liv yo.  
      John bought the book vs. John bought the/their books.

This seems to be quite similar to the plural marker of Papiamentu, which also postpones a plural marker specified for specificity, i.e. *nan*. Both the Haitian plural marker *yo* and the Papiamentu plural marker *nan* are homophonous with the forms of the third plural pronoun.

Despite the lack of nominal plural morphology, HC has the distinction mass vs. count nouns. This is particularly evident if you take a look at the sentence below:

- 489) Mwen manje diri \*yo.  
      I ate rice/\* I ate the rices/ I ate several types of rice

Depréz notes, in fact, that an overt pluralization of *diri* ‘rice’ is possible only if a taxonomic reading must be yielded. Moreover, in order to be “partitioned”, a mass noun must be introduced by a measure phrase (2005:862). This is not different from the case of GBC, as we described in Chapter 4. Finally, it also happens in classifier languages such as Mandarin

Chinese and Vietnamese, as we saw in Chapter 5. These languages use a measure phrase to properly partition a mass noun. On the other hand, they use classifiers in order to refer to (and specify the – singular or plural - number of) the individuals.

So far, we said that HC singular bare nouns, or rather BNs underspecified for number, may yield both singular and plural readings. Depréz explains that there is a certain variation from speaker to speaker, which seems to depend upon both sociolinguistic and pragmatic factors (2005:863). One test to see whether a bare noun yields a singular or plural reading is, as advised by Depréz, to check whether there is an eventual anaphoric pronoun. The number interpretation of the BN will be consonant upon the grammatical number of the pronoun. We will not see these facts in more detail here. For our present discussion, we just need to know that both singular and plural interpretations are available for HC BNs.

HC BNs may yield kind-referring, generic and existential readings. The sentences in (490-492) below, from Depréz (2005), will clarify the facts explained so far:

- 490) *Elefan ap vin ra.*  
‘Elephants are/the elephant is becoming rare.’
- 491) *Jouromou pa donnen kalbas.*  
‘The ‘jouromou’ cannot produce calabashes.’
- 492) *Ou pat we tach até a.*  
‘You did not see spots on the floor.’

The sentence in (490) contains the bare noun *elefan* ‘elephant’, which is the subject of the kind-referring predicate *vin ra* ‘become rare’. It yields, thus, a kind-referring reading. *Jouromou* and *kalbas* ‘calabashes’ in (491), in the subject and object positions respectively, are object-referring BNs and receive a nonspecific reading. In other words, they are generic. The sentence in (491), however, is a generic (characterizing) sentence and that is what yields the nonspecific reading of *jouromou* and *kalbas*.<sup>115</sup> Finally, as for existential readings, bare nouns in HC can only have narrow scope readings with respect to negation and intensional verbs. In (492) we have an example of this: *tach* ‘spot’ may only yield a nonspecific reading since it takes a narrow scope with respect to the negation *pat*. As remarked by Depréz, in order to derive a specific (wide-scope) reading with respect to negation, it is necessary that

---

<sup>115</sup> Remember that, as we said in Chapters 3 and 4, specific object-referring NPs may also occur in characterizing sentences: however, this is not the case at stake here.

the indefinite determiner *yon* introduce *tach*; *yon* may have both a wide-scope and narrow-scope reading (493):<sup>116</sup>

- 493) Ou pat we yon tach até a.  
'You did not see one spot on the floor/there is one spot on the floor you didn't see.'

This is different from the case of GBC, where the bare noun always seems to take both wide and narrow scope, also in case of negation.

It is worth noting that HC grammar does not allow definite readings of bare count arguments. In order to license this reading, a definite or demonstrative determiner should preferably follow the bare noun, i.e. *la* or *yo*, singular and plural definite markers, respectively. As we mentioned above, postnominal *yo* is the only number marker, i.e. plural form, that HC has, and it is, however, limited to definite nouns.

The facts described so far lead us, in the spirit of Depréz (2005, 2006, 2007), to the conclusion that Haitian Creole is a –PL language. HC BNs have, indeed, direct access to kind-referring readings. In these cases, NumP does not project. Also, existential and generic readings are predicted which, in these cases, NumP projects. The most striking difference from English, and from +PL languages in general, is that the HC bare noun is underspecified as for number. It can thus yield both singular and plural interpretations. In +PL languages, on the other hand, NumP always projects. As a consequence, bare nouns are always specified for number. In other words, BNs in +PL languages are always either singular or plural. So, also in this case of –PL language, according to Depréz, we saw that there is no plural restriction, thus crucially diverging from +PL languages.

#### 6.4.2.2 Portuguese-based Creoles: CVC and GBC

As remarked in Depréz (2007:314), CVC's behavior with respect to number marking is quite different from Haitian Creole. HC does not, in fact, overtly mark nouns as plurals, at least not without marking them as definite as well. However, CVC can mark plurality on the noun directly, through the plural suffix *-s*. In this regard, CVC patterns together with GBC and their common lexifier European Portuguese (and Romance and Germanic, in general), in

---

<sup>116</sup> The scopal properties described so far with respect to HC BNs do not hold in case of modified nouns: whenever a modifier is present, the noun does not denote a kind (Depréz 2005:860).

having the possibility of marking plurality with a plural suffix directly adjoined to the noun, i.e. bound morphology. CVC (and GBC, as we will see below) differ from EP with respect to Depréz's Plural Parameter. EP is a +PL language, which does not allow bare singular count nouns in the argument position, thus undergoing the above described plural restrictions.

As we already saw in Chapter 5, a noun introduced by quantifiers or numerals is usually not marked as plural. This seems to hold in CVC for nouns introduced by determiners and demonstratives as well. We have to introduce here an important difference: as we already observed in Chapter 4, it is the noun in GBC that usually receives the mark of plurality, and not the (indefinite) determiner, nor the demonstrative. It could be that it is just a tendency or a possibility in one of the Kriyol subgrammars. However, at least in my corpus, this tendency seems to be quite strong. Therefore, with respect to overt (lexical) DPs, we need to take GBC and CVC separately.

As for the number interpretation of BNs in CVC and GBC, this is not different from the HC case: BNs in both Portuguese-based creoles may be interpreted as singular or plural, when they lack overt plural morphology. According to this model, "morphologically unmarked BNs do not have to denote a mass to be acceptable" (315). This means that morphologically nonplural BNs in HC, CVC and GBC, and in creoles in general, are underspecified for number. Therefore, Chierchia's (1998) prediction that they be mass is incorrect. As in the case of GBC, CVC, at least in its basilectal variety, usually marks plurality only with [+animate] nouns.

As a crucial difference from HC, both CVC and GBC happen to have a noun class, whose nouns behave as +PL, namely the noun class [+human, +animate]. Nouns which belong to this class are usually overtly pluralized, whenever more than one entity is implied. The situation seems to be more complex in GBC since it also involves referentiality. We already spoke about these facts in Chapters 3 and 4, and therefore omit them from the present discussion.



## 6.5 An alternative approach to Depréz’s proposal: Introductory remarks

The alternative model proposed in the present section is inspired by Depréz’s (2005, 2006, 2007) approach, namely her Plural Parameter. In the following discussion we are not going to revisit the syntactic perspective discussed in the work of Depréz. According to her, we assume that bare nouns, both  $\neg$ [number] and plural, are DPs. More specifically, they project at NumP, whenever they have morphological overt number, i.e. either singular or plural in +PL languages, overt plurals in –PL languages, or with an overt (in)definite determiner or a numeral in both groups of languages. Therefore, the only case they project at NP are  $\neg$ [number] bare nouns. Thus, it happens in –PL languages only. In any case, at D-structure, bare nouns are always full DPs. We may retain this syntactic perspective in the spirit of Abney (1987) and, as for bare nouns in particular, Depréz (2005, 2006, 2007) among others. The present discussion will not deal with this topic in greater depth. For the time being, our purpose is to develop a model able to predict the ways of forming Number Phrases crosslinguistically, to motivate the variation from a theoretical perspective, and to build a linguistic taxonomy with respect to the category number.

### 6.5.1 An alternative account

We will discuss in the present section the derivation of (sets of) object individuals, i.e. specimens, from kind individuals. We assume with Depréz (2007) that the base denotation of nouns (i.e. the denotation of the morphological root of the noun) crosslinguistically belongs to the type e of individuals. In (494), the kind ‘dog’ is represented by the Latin name for dog, i.e. *Canis*, and the kind ‘water’ is named after his chemical composition, i.e.  $H_2O$ :

494)	‘Dog’:	kind <i>Canis</i> (C)	
	‘Water’:	kind $H_2O$	
	a) [NP dog]	type e	C
	b) [NP water]	type e	$H_2O$

In the representations in (494a) and (494b), N(oun) Phrase could be, in principle, replaced by Name Phrase. Here, in fact, NPs behave like proper names (cf. Carlson 1977 for linguistic evidence that NPs that refer to kinds are names). Following Krifka et al. (1995), “kind-referring NPs refer to well-established entities in the background knowledge of speaker and

hearer” (1995:69). Moreover, on the basis of the fact that they denote kinds, which are well established in the background knowledge too, Krifka et al. claim that they are lexical entries and cannot be analyzed as complex expressions. In other words, their meaning cannot be derived from their parts. They should, thus, be considered as idiomatic expressions, and, in this sense, kind-referring NPs behave like proper names. We are not going to replace NP by Name Phrase in order to keep our derivational operation close to the syntactic positions into which they map.

In order to derive objects from kinds, as we already saw in Depréz’s proposal, Carlson’s realization rule R applies to kinds. In principle, we may say that this takes place in NumP; however, since we are going to obtain a set by means of this operation, a more descriptive name for NumP at this stage could be SetP. With the operation R, in fact, we are not deriving a semantic number, i.e. singular or plural. This stage will be achieved with the next step, as we will see below. What we obtain by means of the application of R is a set, which can be composed of individuals or sums of individuals. For instance, if we want to derive the set of dogs from the kind *Canis*, the realization operation R will apply to the kind *Canis*. This is shown by means of lambda-abstraction in (495a). At this point, we should recall the difference between count and mass nouns that we introduced in Chapter 4 with respect to GBC BNs. *Dog* is a count noun referring to an atomic entity whereas *water* is a mass noun, which does not refer to any atomic entity, unless it yields a taxonomic reading.

We are able now to account for object-referring readings of mass nouns by means of the same operation as we saw in the case of count nouns. This is shown in (495b), where we derive the set of quantities of the kind *H<sub>2</sub>O* in the same fashion as for the count noun *dog* in (495a), namely by means of application of R. The crucial difference is, thus, that for the count noun in (495a) we derive the set of dogs, either individuals or sums. On the other hand, in (495b) we yield the set of quantities of the kind *H<sub>2</sub>O*.

- 495) a.  $[\text{SetP } \textit{dog}]$ ; type et;  $\lambda x[\text{R}(x, \textit{C})]$ ;  
 b.  $[\text{SetP } \textit{water}]$ ; type et;  $\lambda x[\text{R}(x, \textit{H}_2\text{O})]$ .

We come now to the stage where number is realized, which will not be different from Depréz’s proposal. Carlson’s realization rule R and Krifka’s Object Unit OU will apply to kinds to yield a function from a number to a set. This is shown in (496) below, where we have the representation of a function from number to sets:

- 496)  $[\text{NumP } \textit{dog}]$  type  $\langle n, \langle e, t \rangle \rangle$   $\lambda n \lambda x[\text{R}(x, \textit{C}) \wedge \text{OU}(\textit{C})(x) = n]$

In the representation in (496) above, we have a derivation for number in case of count nouns. There is, however, another reading to consider with respect to kinds, namely the taxonomic reading. Following Krifka et al. (1995), the taxonomic reading of *dog* will be yielded by means of the application of the taxonomic rule T and the Kind Unit KU. This is represented in (497):

$$497) \quad \lambda n \lambda x [T(x, C) \wedge KU(C)(x) = n]$$

We return now to the step where we get a number for object-referring readings, namely the function from number to sets. We saw in (497) how to derive a number realization for count nouns such as ‘dog’. This is, however, not possible for mass nouns such as ‘water’, which is not unexpected, if we consider that object-referring mass nouns cannot have plural readings like count nouns. As for referential nouns, plurality implies, in fact, that more than one entity is referred to. There seems to be, however, an exceptional case, namely decimal numbers like 0.8 or 1.0. They always require a plural noun in English (Krifka p.c.). This holds also in case of 1.0, which translates into ‘a unity’. The same holds for Italian and probably many other languages, but we will not examine this point in more detail since this is not crucial to the present discussion. Unlike count nouns, as we mentioned above, mass nouns do not have atomic readings. As a consequence, they cannot be pluralized for yielding more than one atomic entity. Therefore, derivations like the one in (498) below will be ill-formed:

$$498) \quad *[\text{NumP water}]; \text{OU}(\text{H}_2\text{O})(x) \text{ is undefined.}$$

So far, we have derived (sets of) objects and numbers of (sets of) objects from kinds. In other words, we have represented the path from kinds of type  $e$  to sets of type  $\langle e, t \rangle$  and to numbers of sets of type  $\langle n, \langle e, t \rangle \rangle$ . In the following section, we will see how these facts apply to different types of languages.

## 6.6 An overview of the features pertaining to different types of languages with respect to the Num requirement

As we saw in Depréz’s proposal and as noted by Depréz herself, the terms ‘+PL’ and ‘–PL’ languages may be misleading if we group the languages following a ‘plural’ requirement, which is not the crucial one for distinguishing between the two groups of languages. As proposed in Depréz (2007), what is critical for the parameter is the projection

of NumP. On this basis, we could therefore choose the terminology +Num and –Num languages, which seems to better describe the point of divergence of the two sets of languages. The NumP projection and the saturation of its semantic argument are, in fact, not only responsible for overt plurality. Crucially, they are also responsible for singularity. This is, however, the situation of the +PL group. In the other group, i.e. –PL, whenever NumP fails to project, we do not yield a singular, i.e. a set which corresponds to one entity or, in other words, a set which contains all the properties of being that particular entity. What we obtain is a bare noun, which is underspecified for any number, i.e. –[number]. On the basis of what has been stated so far, we will propose the two macro-groups of languages –Num and +Num. The change is, as we will see in the course of the following discussion, not only relative to the naming of the parameter. Our taxonomy will, in fact, include a further branching of both macrogroups, and, more specifically, –Num will contain languages with classifiers, on the one hand, and without classifiers, on the other. The +Num group splits into languages that require number agreement between the number word and the noun, and languages that do not have such a requirement. We will see the facts in more detail below.

We will see now in more detail the setting of both –Num and +Num language groups. As we will see, the most striking difference between the two groups is, as noted in the principle underlying Depréz's proposal, the optionality of the projection of NumP: +Num languages always project NumP for count nouns and saturate the number argument. On the other hand, –Num languages allow for the optionality of NumP projection for their count nouns. As for mass nouns, the two groups do not differ from each other: NumP does not project, and the mass noun is yielded in SetP. In languages such as Guinea-Bissau Creole, this could easily translate into the situation described in Kihm (1994, 2007) as for overt plurality. In GBC, number specification seems to occur whenever it is relevant in the discourse world shared by speaker and hearer. In such cases, as already described in Kihm and as we saw in Chapter 4, animacy and referentiality play a crucial role, at least in one of the subgrammars available to GBC native speakers. We already saw that bare nouns in such languages are underspecified for number. Recall that this means that a bare noun –[number] may yield both singular, plural and number neutral readings. The presence of anaphoric pronouns, singular or plural, will help us derive the correct interpretation in such cases.

The languages belonging to the –Num group allow for SetP in case of both mass and count nouns. On the other hand, they allow for NumP as an option in case of count nouns, only. These facts were shown above in the representations in (495) and (496), which we re-propose here in (499a, b and c), respectively:

- 499) a. [SetP dog]; type et;  $\lambda x[R(x,C)];$   
 b. [SetP water]; type et;  $\lambda x[R(x,H_2O)];$   
 c. [NumP dog] type n et  $\lambda n\lambda x[R(x,C)\wedge OU(C)(x) = n].$

As for +Num languages, they pattern together with the –Num group with respect to SetP, which in case of +Num also allow for both mass and count nouns. The point of divergence is, thus, straightforwardly derived. In contrast to –Num, +Num have to obligatorily satisfy the number requirement and, thus, overtly express singularity and plurality. Otherwise the noun will result uninterpretable. Therefore, +Num languages always require NumP for count nouns. The situation described so far is what underlies the proposal of Depréz. Just as in Depréz’s model, this very same point is the trigger for our language taxonomy.

We come now to the setting of languages such as Chinese and Vietnamese, namely classifier languages. They allow for SetP for both mass and count nouns. This opposition is made explicit by means of classifier. We saw these facts in Chapter 5 with respect to Mandarin Chinese and Vietnamese, respectively. As a difference from both the –Num and the +Num groups, NumP in classifier languages do not have implicit OU/KU operations. The derivations in (496) and (497) above are, thus, not applicable to the case of classifier languages. This point has to be further developed, but for the time being, we assume that classifiers are used in order to yield number readings and taxonomic readings in these languages.

Before we propose a summary of our findings, we should look at measure phrases. It holds for all languages that measure phrases are applied to set phrases, i.e. SetP. In other words, measure phrases do not yield number specification. Therefore, unlike items directly responsible for number such as overt plural markers, number words and (in)definite determiners, measure phrases cannot satisfy the number argument of the semantic counter. The findings described so far are summarized in Table (11) below.

Table 11. *SetP and NumP in different types of languages*

	SetP	NumP
-Num	Yes	No/Yes
+Num	Yes	Yes
+Cl	Yes	Yes

In Table (11) we call +Cl those languages, which use classifiers to express number. As we already mentioned, they will belong to the -Num group of our language taxonomy. We will see these facts in more detail in the following discussion. For the time being, we simply aim at looking at the setting of these languages.

In Table (12) the facts about the number requirement are summarized. The number argument is strictly necessary for count nouns in +Num languages. Both the -Num group and its cluster +Cl show optionality, which indicates the fact that both groups of languages allow for bare nouns unspecified for number. This is not the case of +Num languages.

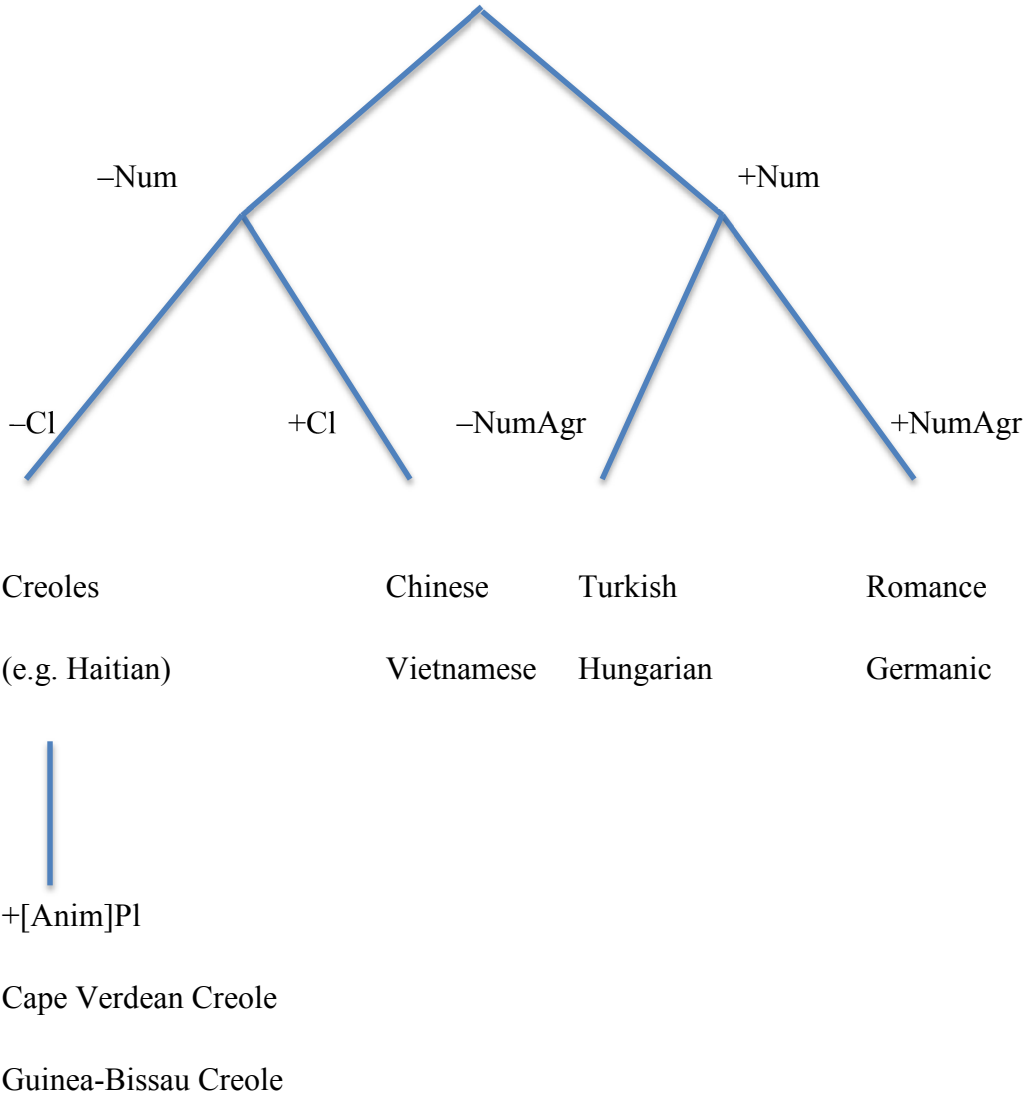
Table 12. *Number requirement in different types of languages*

	¬[number]	[+number]
-Num	Yes	Yes
+Num	No	Yes
+Cl	Yes	Yes

### 6.7 Predictions based on the above discussion relative to each group

Before we enter the heart of the matter and describe the characteristics of the different sets of languages with regard to number as predicted by our proposal, it could be useful to have a look at Figure 5 below. Here, we have a visual representation of the different sets of languages and their branching based on their behavior with regard to number.

Figure 7 *Taxonomy of languages based on number*



We will discuss the linguistic branching represented in Figure 7 in more detail in the following sections. For the time being, however, it is useful to have a visual representation of what we will say below.

As we said above, the basic denotation of nouns in all languages is kind. This reading, as noted in Depréz (2005, 2007), is directly accessible for –PL languages. In the present alternative proposal, we will say that the kind-referring reading is directly available for –Num and +Cl languages, and for mass nouns in +Num languages. This is shown below in (500). More specifically, in (500a) we have the representation of the kind-referring reading for the count noun *dog* in –Num and +Cl languages. The kind-referring reading is, in the case at hand, made possible by the predicate ‘be extinct’, which typically occurs in relation to kinds. The same can be said as for the example in (500b). However, unlike (500a), the representation in (500b) concerns mass nouns in all types of languages, namely –Num, +Cl and + Num. For both mass and count nouns considered in this point, it holds that it happens at NP:

- 500) a. [NP Dog] is extinct; EXTINCT (C);  
 b. [NP Water] is abundant; ABUNDANT (H<sub>2</sub>O).

As for the existential reading for both mass and count nouns in –Num and +Cl languages, and for mass nouns only in +Num languages, the existential reading arises at SetP. Existential closure is needed. In (501a) we have the bare count noun ‘dog’ in –Num and +Cl languages in an existential sentence, and (501b) shows how to derive the existential reading by means of existential closure. In a similar fashion, the existential reading of the bare mass noun ‘water’ in (501c) will be explained via existential closure as shown in (501d). Recall that ‘water’ in (501d) represents any mass noun in all languages. The crucial difference between (501b) and (501d) is that the ‘dog’ (y) in (501b) is either atomic or represents a sum of atomic individuals, whereas ‘water’ (y) in (501d) does not imply any atomic or sum distinction. As a consequence, the count noun ‘dog’ in (501a), which in the present case is underspecified for number, may be referred to by an anaphoric pronoun, either singular or plural. As expected, the mass noun ‘water’ in (501c) may be anaphorically bound by a singular pronoun only. A plural one is not allowed since it presupposes atomic proper parts.

- 501) a. [SetP Dog] COP barking, because it/they is/are hungry;  
 b.  $\exists y[\lambda x[R(x,C)](y) \wedge \text{BARKING}(y)]$   
 $\Leftrightarrow \exists y[R(y,C) \wedge \text{BARKING}(y)]$   
 c. [SetP Water] is dripping;



- d.  $\exists y[\lambda x[R(x,H_2O)](y) \wedge \text{DRIPPING}(y)]$   
 $\Leftrightarrow \exists y[R(y,H_2O) \wedge \text{DRIPPING}(y)]$

Now we come to the case of existential readings of count nouns in –Num and +Num languages. As we saw in the discussion above, in order to get an interpretable number argument, we need OU, which, therefore, must be defined. This is, of course, not possible for mass nouns. As we can see in (502a), ‘one dog’ projects at DP, which holds for both –Num and +Num languages. The case of +CL languages, on the other hand, will be dealt with separately. The derivation is represented in (502b), where existential closure binds the DP:

- 502) a.  $[\text{DP one}_{[\text{NumP dog}]}] \text{ COP barking};$   
 b.  $\exists y[\lambda n\lambda x[R(x,C) \wedge \text{OU}(C)(x) = n](1)(y) \wedge \text{BARKING}(y)]$   
 $\Leftrightarrow \exists y[R(y,C) \wedge \text{OU}(C)(y) = 1 \wedge \text{BARKING}(y)]$

The case of numerals offers another interesting example and, importantly, introduces us to a further branching of the +Num group. There are languages that do not require number agreement between numerals and nouns, which implies that whenever a number word different from ‘one’ introduces a noun, the noun will be not marked for plurality. This seems to be the same case as creoles, or at least the creoles treated so far, which belong to the –Num group. Yet, Turkish behaves rather as a +Num language: its bare plurals trigger, in fact, a number reading where the entities at stake are necessarily more than one. The fact that English bare plurals are number neutral will lead us to the subdivision within the +Num group, as we mentioned above, and we will see this in more detail below. For the time being, we have to illustrate the formation of numeral phrases in languages such as Turkish. Similar to the former case, we derive in (503b) the meaning of the numeral phrase ‘two dog- $\emptyset$ ’ (represented in 503a), which could represent, in principle, the case of Turkish or of any other language of the type [+Num, –NumAgr]. An example of Turkish number phrase is as follows: *kırk harami* ‘forty thief’.

- 503) a.  $[\text{DP two} [\text{dog}]] \text{ COP barking};$   
 b.  $\exists y[\lambda n\lambda x[R(x,C) \wedge \text{OU}(C)(x) = n](2)(y) \wedge \text{BARKING}(y)]$   
 $\Leftrightarrow \exists y[R(y,C) \wedge \text{OU}(C)(y) = 2 \wedge \text{BARKING}(y)]$

We come now to the case of numeral phrases in +Cl languages. In this case, a classifier will be inserted in [Spec, NumP], resulting at the right of the number word, which shows up in DP. This is shown in (504a). In the formula in (504b), the application of R and OU to a kind K gives us a number, which is achieved by means of an operator, i.e. the classifier. Finally, in

(504c) we replace the kind  $K$  with the kind *Canis*, and obtain a function from a number  $n$  to the set of sum individuals  $x$  that are dogs and that have  $n$  members. If we compare (504c) with the derivation in (496) above, we will easily notice that these are identical operations.

- 504) a. [DP one [NumP CL [NP dog]]] is barking.  
 b. CL:  $\lambda k \lambda n \lambda x [R(x, K) \wedge OU(K)(x) = n]$   
 c. CL [dog]:  $\lambda n \lambda x [R(x, C) \wedge OU(C)(x) = n]$

Finally, in +Num languages NumP must be formed whenever possible. This means that NumP must project, as the grammaticality of (505a) will show. On the other hand, it will be uninterpretable (505b):

- 505) a. [NumP dog] COP barking;  
 b. \*[SetP dog] COP barking.

We have described so far the crucial features of different types of languages with respect to number. In the discussion above, we mentioned four groups of languages, namely –Num, +Num, +Cl and –NumAgr. On the basis of the differences observed so far, we will build our linguistic taxonomy with respect to issues pertaining to the grammatical category Number. Before we get to this final point, we will review the different ways languages have to form Number Phrases. We will, consequently, extend our discussion to two further groups of languages and will be, finally, able to elaborate a complete taxonomy based on ‘number’.

## 6.8 Ways of forming NumPs

We have essentially two ways of forming NumPs or rather we have different morphological ways of forming NumPs, which go back to two semantic representations. This will lead us to the (initial) binary branching of our language taxonomy. We will return to this point later but, for the time being, we need to look in more detail at the strategies that different languages have to form NumPs.

Above, we already saw the way –Num languages form NumPs. More specifically, they project NumP whenever an element involving more than one individual is present. It could be a numeral, overt plural morphology or an overt determiner. The representation is the

same as in (502), re-proposed here in (506) with a slight change: ‘one’ has been substituted by ‘two’ to obtain a plural interpretation. The case at stake here is what we saw for creoles.

- 506) a. [DP two<sub>[NumP dog]] COP barking;  
 b.  $\exists y[\lambda n\lambda x[R(x,C) \wedge OU(C)(x) = n](2)(y) \wedge BARKING(y)]$   
 $\Leftrightarrow \exists y[R(y,C) \wedge OU(C)(y) = 2 \wedge BARKING(y)]$</sub>

The same holds for languages such as Turkish and Hungarian, which do not show agreement between number words and the nouns they modify. However, such languages, say –NumAgr languages, do not pattern on a par with –Num languages such as creoles. Recall, in fact, that languages belonging to the –Num group allow for bare nouns unspecified for number. Crucially, this is not possible in –NumAgr languages.

A different way of forming NumPs is represented by languages such as English, where the distinction between singular and plural matters. This is reflected in the morphology of number in English. We generally have, in fact, morphologically singular forms different from morphologically plural ones. This opposition is represented in (507a) and (507b): starting from ‘dog’ in (507), we derive the form ‘dog’ with singular specification in (507a) and the form ‘dogs’, specified for plural in (507b). In its initial stage in (507), ‘dog’ is of the type e. If OU did not apply, the result would be a SetP, which is not specified for number and, as a consequence, is not interpretable. If OU co-applies with R, the result is a well-formed NumP, and the NumP will be interpretable, as shown in (507a) and (507b).

- 507) [NP dog]  
 a. sg: [NumP dog]:  $\lambda n\lambda x[R(x,C) \wedge OU(C)(x) = n \wedge n=1]$ ;  
 b. pl: [NumP dogs]:  $\lambda n\lambda x[R(x,C) \wedge OU(C)(x) = n$  (possibly:  $n > 1$ )]

In (507a), we obtain a number argument  $n$ , which must be specified. Provided that singular means ‘one entity’, the only value available for  $n$  is 1. On the other hand, in (507b)  $n$  represents plurality, which has the following requirement: possibly  $n > 1$  or, in other words, the number argument should be possibly higher than 1.

As for bare plurals, the number argument is existentially bound in English, as shown in (508). The formula in (508c) represents the argument taken by -s, where  $N$  is the semantic counterpart of  $n$ . Finally, (508d) represents the existential closure applied to the number argument and the requirement that  $n > 1$  (which translates into –s), on the one hand, and to number of sets of objects and the requirement that  $n > 1$  (say ‘dogs’), on the other.

- 508) a. [DP dogs] COP barking.  
 b. [DP [NumP dog] -s];  
 c. -s:  $\lambda N \lambda x \exists n [N(n)(x) \wedge n > 1]$   
 d. Dog -s:  $\lambda N \lambda x \exists n [N(n)(x) \wedge n > 1] (\lambda n \lambda x [R(x,C) \wedge OU(C)(x) = n$   
 $= \lambda x \exists n [R(x,C) \wedge OU(C)(x) = n \wedge n > 1]$

Notice that in English *n* is not necessarily higher than 1. As we already mentioned above, *n* is unspecified in case of bare plurals.

Now we come to a different way of plural marking: there are, in fact, languages such as the creole language Bislama, which form the plural by means of pre- or postnominal insertion of an item homophonous with the third plural pronoun. Postnominal insertion of this item seems, however, to be either the Ambrym variety of Bislama or a possibility in Bislama, in general (Krifka, p.c.). The facts explained above with respect to the plural formation in Bislama via postnominal insertion are represented in (509). In (509d), where *P* is a variable of type  $\langle e,t \rangle$ , *x* has more than one atomic part, i.e.  $AT(x) > 1$ . Finally, (509e) is the representation of ‘dog they’. This seems to be similar to the case described in Chapter 5 as for plurality in Santome and Papiamentu. Both creoles express overt plurality by means of a plural marker homophonous with the third plural person. In more detail, the plural marker *nan* in Papiamentu is postnominal, which seems to parallel the case of Bislama. On the other hand, the prenominal *inen* in Santome is better described as a plural definite determiner based on the fact that whenever it introduces a noun, the noun yields a plural definite reading (Alexandre and Hagemeyer 2007). Both cases could be a development from something like ‘dog they’ in (508a). The case of Santome, however, would need a further change into a prenominal determiner position.

- 509) a. sg: dog it vs. pl: dog they  
 b. [DP dog they]  
 c. dog: [SetP dog];  $\lambda x [R(x,C)]$   
 d. they:  $\lambda P \lambda x [P(x) \wedge AT(x) > 1]$ ;  
 e.  $\lambda P \lambda x [P(x) \wedge AT(x) > 1] (\lambda x [R(x,C)]) = \lambda x [R(x,C) \wedge AT(x) > 1]$

Interestingly, whenever a number word introduces the noun, no plural marker is required, as shown by the ungrammaticality of the example in (510):

- 510) \*three [dog they]

In (511) we have examples from Bislama and its way of forming NumPs. We will see here examples of overt plurality via prenominal insertion. In the context of our analysis, it does not make any difference if the plural item is inserted before or after the noun. As we can see from the examples below, there is no agreement between the number word and the noun. As a consequence, (511d) is ungrammatical. It is interesting to notice that Bislama has two third plural pronouns, i.e. *ol* and *olgeta*. Both may be used as plural markers. Some speakers use *ol* for animates and *olgeta* for inanimates. *Ol* is the clitic pronoun, whereas *olgeta* is the full pronoun.

- 511) a. wan dog 'one dog'  
 b. tri dog 'three dogs'  
 c. ol dog 'dogs'  
 d. \*ten olgeta dog 'ten dogs'

These would translate into the representations in (512), where (512c) represents the sum of all dogs, provided that there is more than one dog.

- 512) a. [DP [SetP dog] they]  
 b. they:  $\lambda P \sigma x [P(x) \wedge AT(x) > 1]$   
 c. dog they:  $\sigma x [R(x,C) AT(x) > 1]$   
 d. \*[DP [DP two [NumP dog]] they]

## 6.9 Taxonomy

In the present section, we will represent our linguistic taxonomy on the basis of the features dealt with so far. The languages are divided into two macrogroups, which we call –Num and +Num. This binary choice is reminiscent of Depréz's proposal. The –Num group, in fact, basically distinguishes itself from the +Num languages in its optionality in NumP projection. As we already saw, this translates into a difference in number specification for bare nouns (without plural morphology). In –Num languages, (nonplural) bare nouns are unspecified for number. On the other hand, such bare nouns in +PL languages are uninterpretable and, thus, not allowed.

We further noticed that there are important differences among languages belonging to each group, which brings us to a further branching of each group. As a consequence, the –

Num group splits into two clusters, which we will call –Cl and +Cl. This is based on the crucial presence of classifiers (operators for number) in the +Cl subgroup. On the other hand, the languages that belong to the –Cl cluster do not have classifiers. Languages such as Mandarin Chinese and Vietnamese belong to the +Cl group while Haitian Creole, Cape Verdean Creole and Guinea-Bissau belong to the –Cl language cluster. Furthermore, recall from the discussion on Depréz’s model that CVC and GBC occur to have a noun class whose nouns behave as +Num (or as +Pl, in Depréz’s terminology). The nouns in this class happen to have the feature [+human, +animate]. As a consequence, we have a further split within the –Cl subgroup of the –Num language group, namely the group of languages that has such gradient of animacy. We could call this group +Anim. In this subgroup, we may add Santome which, as we noticed in Chapter 5, may mark plurality (always together with definiteness) by means of the plural definite marker *inen* in case of nonmodified nouns which are [+human]. This does not seem to happen in French-based creoles (Depréz 2005, 2007).

We should at this point notice that the animacy gradient plays an important role in the nominal systems of (at least some) noncreole languages as well (Corbett 2000), e.g. the same gradient of animacy is said to act in the nominal system of Vietnamese (Nguyen 2008). The consequence of animacy in this +Cl language, as we already observed in Chapter 5, is reflected in the classifier system. Vietnamese has, in fact, a specialized classifier for each noun class [+human], [+animate] and [–animate]. This leads us to leave this issue open and consider that, in principle, each group of the present taxonomy could internally distinguish between languages with animacy requirements, on the one hand, and languages which do not have specialized noun classes for animacy, on the other, which could be the object of future study. For the time being, we just take note of the existence of a further subgroup within the [–Cl, –Num] branch of our taxonomy, say +Anim to indicate that animacy (at least +human) matters.

Now we come back to the initial binary branching and discuss the +Num group. Recall that the languages of this group always project NumP and, thus, satisfy the number requirement. The branching of +Num gives rise to the two following subgroups: +NumAgr and –NumAgr. The +NumAgr cluster contains languages such as Romance and Germanic languages. As we discussed in the case of English, these languages always require agreement between the number word and the noun. Moreover, as for English we said that bare plurals are number neutral, which is why they need to be existentially bound. However, languages such as Turkish do not have this requirement. Their bare plurals are specified for plurality; in other words, plurality in Turkish translates as more than one. Nonetheless, it is important to note

that both languages have bare plurals. Moreover, numeral words in –NumAgr languages are followed by a nonplural noun. This seems to be the case of the creoles analyzed in the present study as well. Yet, we would better classify these creoles together with the other creoles described in Depréz, namely French-based creoles. This choice is based on the fact that both Santome and Papiamentu have bare (nonplural) nouns, which are unspecified for number and can, thus, be interpreted as both singular and plural.

In Figure 8, we have a ‘visual’ representation of the facts described so far. Figure 6 corresponds to Figure 7 above. Concluding, the NumP/SetP proposal discussed here has, thus, a twofold theoretical strength. Starting from the same principle as Depréz’s proposal, it has the force of explaining the bare singular/bare plural asymmetry between creoles, on the one hand, and Romance and Germanic, on the other. Finally, it predicts a very straightforward taxonomy of languages with respect to the realization of grammatical number. The representation of our taxonomy is in Figure 8.<sup>117</sup> Here other languages have been inserted on the basis of the above discussion on the category ‘number’ and its realization in different languages. For instance, Gbe languages belong to the same group as creoles, i.e. [-Num, -Cl] since their BNs behave very similar to creole BNs. Moreover, Gbe languages do not usually show agreement between the number word and the noun. Nonetheless, as we stated in Chapter 5, Gbe languages use plural morphology as a strategy for expressing specificity.

---

<sup>117</sup> Languages such as Tariana do not seem to fit any group of the typology in Figure 8. Tariana has both classifier and agreement between the number word and the noun. Such a situation is not dealt with in our present typology of languages. There is need for further studies in this direction. It is, however, interesting to note that the plural marker in Tariana does not usually occur on inanimate nouns modified by a numeral. This is a strategy for individualization (Krasnoukhova 2012:103). Thanks to Pieter Muysken for making me aware of this (p.c.).

E.g. Tariana (Arawakan; Aikhenvald 2003:217)  
 Kephunipe-phi-pe surupe-phi-pe.  
 four-CLF :hollow-PL clay-CLF :hollow-PL  
 ‘Four clay pots’.





## CHAPTER 7

### Conclusions

#### 7.1 Summary and generalizations

The first topic of the present dissertation is the description of bare nouns in Guinea-Bissau Creole. Both distribution and interpretation are taken into account. Furthermore, this description is inserted in the broader perspective of the study of the whole nominal and determiner system of GBC.

These topics are dealt with in Chapter 4, where we also give a more restricted definition of Kriyol bare nouns (and BNs in creoles in general). They are noun phrases without any overt determiner and without any specification for number. Bare nouns in GBC may occur in both the subject and object positions. On the one hand, bare subjects show a strong preference for the definite interpretation. The left-periphery of the sentence is, in fact, associated with ‘presupposedness’. As a consequence, what is already known, e.g. what has already been mentioned in the discourse world shared by speaker and hearer, is more likely to find expressions in leftward positions, i.e. topics and subjects. Similar to GBC bare subjects, also recipient objects yield a definite reading. As a difference from both subject and recipient, patient objects may yield any possible reading, except the plural specific.

In our description of the Kriyol nominal system, we noted that Kriyol has no definite determiner. The expression of definiteness principally relies on the noun only. In other words, bare nouns express definiteness, at least in the case of bare subjects. Just in certain cases such as anaphoric and determinate reference, we may find a noun introduced by the distal demonstrative *kil* ‘that’. The use of *kil* is, however, not mandatory. It is simply an option in the mind of the speaker. As for indefiniteness, it may be expressed via BNs or via the indefinite determiner *un* ‘a/an’. Importantly, the indefinite determiner *un* seems to be semantically associated with nonspecificity. On the other hand, indefinite BNS are unspecified as for specificity. There seems, thus, to be in GBC a tendency to overtly mark the contrast specificity vs. nonspecificity more than the contrast newness vs. presupposedness. As for plurality, we saw that a noun is overtly marked as plural when its number specification is

relevant to the discourse context. Further conditions for overt plurality are animacy and referentiality.

Furthermore, we saw that BNs in Kriyol are unspecified for semantic number. They may show up in the argument, non-argument and predicate positions. From our description of BNs in Chapter 4, it turned out that there is no subject/object asymmetry as for BN distribution. As for their interpretation, the situation is different: BNs show a certain asymmetry. Bare subjects show a strong preference for the definite reading. The same can be said as for the recipient object, which always yields a definite reading. On the other hand, no reading seems to be banned as for patient BNs, except for the plural specific one. It is in fact quite unlikely that a BN yields such a reading. The overt plural would fit better.

Finally, we discussed the factors that play a role in the derivation of the correct interpretation of GBC BNs. They are aspect and tense as for bare objects, on the one hand, and predicate types as for bare subjects, on the other. Habitual aspect triggers indefinite readings for object BNs. As a difference, perfective and continuous aspects trigger definite readings. For bare subjects, individual-level predicates trigger generic interpretations, whereas stage-level predicates trigger existential (in)definite readings.

The present study also aims at a crosslinguistic perspective among and beyond creoles. First, we undertook a crosslinguistic comparison among languages (mostly creoles), which share Portuguese as their lexifier. In the case of Brazilian Portuguese, it is a variety of EP spoken in Brazil, which is often described as a semi-creole. Furthermore, Papiamentu also has Spanish as its lexifier. The situation with respect to lexifier language(s) in Papiamentu is not simple and has been described in more detail in Chapter 5, section 5.3. The second part of the comparison involves noncreole languages. More specifically, we have studied the cases of Mandarin Chinese, Vietnamese and Gbe languages (with particular emphasis on Gungbe and Fongbe). Although these languages, both the creoles and the noncreoles, are typologically distant languages, they share very similar tendencies. All the languages dealt with in Chapter 5, in fact, allow bare nouns in both the subject and object positions. Moreover, bare nouns may yield both singular and plural readings. Although the languages of the noncreole group have further lexical means to express definiteness and/or specificity, bare nouns in both creoles and noncreoles share the same range of interpretations.

In the creole group, the general situation is as follows: bare nouns may occupy both the subject and object position, and may yield both singular and plural interpretations, kind-referring, generic and existential readings. The situation in Papiamentu and Brazilian

Portuguese is slightly different. More specifically, singular bare nouns in Papiamentu may yield both singular and plural readings, at least in the object position. On the other hand, plural bare nouns are more restricted. Bare subjects may occur only in episodic sentences. As a consequence, generic readings are not allowed. Bare plurals in the object position are allowed only if modified. As for Brazilian Portuguese, bare singulars may yield kind-referring, generic and existential interpretations. They may have both singular and plural interpretation, but only with individual-level predicates. They may also show up in episodic contexts. Plural bare subjects may occur in both generic and episodic contexts. Bare plurals may also occur in the object position. Since it is always difficult to speak about commonalities and differences without a visual support, I will re-propose here the tables for the summary of the findings which we saw in Chapter 5. In Table 13, we have a representation of the distribution of BNs in the creole group. On the other hand, Table 14 illustrates the possibilities of interpretation.

Table 13. *Distribution of BNs in creoles*

	Subject: Bare SG	Subject: Bare PL	Object: Bare SG	Object: Bare PL
GBC	Yes	Yes	Yes	Yes
CVC	Yes	Yes	Yes	Yes
Santome	Yes	No	Yes	No
Papiamentu	Yes	Yes/No	Yes	Yes/No
BP	Yes	Yes	Yes	Yes

Table 14. *Interpretation of BNs in creoles*

	Generic	Existential	Definite	Indefinite	Specific	Nonspecific
GBC	Yes	Yes	Yes	Yes	Yes	Yes
CVC	Yes	Yes	Yes	Yes	Yes	Yes
Santomé	Yes	Yes	Yes	Yes	Yes	Yes/No
Papiamentu	Yes/No	Yes	Yes	Yes	Yes	Yes/No
BP	Yes	Yes	Yes	Yes	Yes	Yes

As for the noncreole group, the situation is quite homogeneous. In Mandarin Chinese, Vietnamese and Gbe languages, bare nouns are unspecified for number. As a consequence, bare nouns may be interpreted as both singular and plural. Furthermore, bare nouns may occur in both the subject and object positions. The possibilities of interpretation as for the noncreoles are basically the same. The exception is the specific interpretation, which is not possible in Fongbe. Specificity in Fongbe is expressed by means of overt demonstrative/specificity markers. Table 15 is a summary of the findings as for the possibilities of interpretation in the noncreole group.

Table 15. *Interpretation of BNs in noncreoles*

	Preverbal Position	Postverbal Position
Mandarin Chinese	Definite; generic; (non)specific	(In)definite; generic; (non)specific
Vietnamese	(In)definite; generic; (non)specific	(In)definite; generic; (non)specific
Gungbe	(In)definite; generic; (non)specific	(In)definite; generic; (non)specific
Fongbe	(In)definite; generic; nonspecific	(In)definite; generic; nonspecific

With respect to morphologically overt plural formation, it is worth noting that the gradient seems to be crucial in the creole group only and, more specifically, in the creoles of Guinea-Bissau, Cape Verde, and São Tomé.

Finally, in Chapter 6 we propose a model and the consequent typology as for number crosslinguistically. This proposal is inspired by Depréz's Plural Parameter. First, we described Depréz' (2006, 2007) model, which distinguishes between –PL and +PL languages on the basis of the projection of NumP. While for +PL languages the projection of NumP is always mandatory, for –PL languages it is optional. The most direct consequence of this fact is that +PL languages do not allow singular bare nouns in the argument positions. –PL languages, on the other hand, do allow singular bare nouns. Moreover, such bare nouns are not real singulars. They are rather unspecified for number since NumP does not project. NumP contains, in fact, a semantic counter which is responsible for number saturation. Whenever NumP does not project, the number argument is not satisfied and the resulting bare noun projects at NP.

We retain from Depréz's account the above described generalizations. We also share her assumption that nouns crosslinguistically are kinds of type *e* and that bare nouns are full DPs. The first difference in our model is the terminology. Since the point of divergence between the two groups of languages lies in the number requirement and not in the plurality, we decided to rename the two language groups as +Num and –Num. The fact that the terms +PL and –PL are not the best candidates was already noted by Depréz (2006).

A further difference in our proposal is that we inserted a SetP (Set Phrase). Here, bare nouns are unspecified for number, but nonetheless interpretable in –Num languages (e.g. bare nouns in creoles). This is a consequence of the fact that Carlson's R (Realization rule) applies in SetP in order to yield a (set of) object(s) from a kind. This holds for both mass and count nouns in –Num languages, and for mass nouns in +Num languages. Mass nouns do not require number specification. On the other hand, a bare count noun in +Num languages cannot be in SetP, but must necessarily project in NumP in order to be interpretable, as we already saw in Depréz's account. In NumP, Krifka's OU (Object Unit) applies together with R. The result is that the number argument is satisfied and bare nouns are interpretable. The formulas below will summarize these findings:

- 513) 1. [<sub>NP</sub> dog]                    type e                    C kind *Canis*
2.     a. [<sub>SetP</sub> dog];    type et;                     $\lambda x[R(x,C)];$                     (count)
- b. [<sub>NumP</sub> dog]    type  $\langle n, \langle e, t \rangle \rangle$                      $\lambda n \lambda x [R(x,C) \wedge OU(C)(x) = n];$  (count)
3.     a. [<sub>SetP</sub> water]; type et;                     $\lambda x[R(x,H_2O)].$                     (mass)
- b. \* [<sub>NumP</sub> water];                     $OU(H_2O)(x)$  is undefined.                    (mass)

Crosslinguistically, there are basically two ways to derive NumPs: one for –Num (both +Cl and –Cl) and –NumAgr languages (which belong to the +Num group, but do not require agreement with the number word). This is represented in (514). The second way is the one used in +Num languages (e.g. English), which always require number specification. The formula in (515a) show the case of singulars, and (515b) represents the case of plural formation, where plural is possibly ‘more than one’.

- 514) a.     [<sub>DP</sub> two [<sub>NumP</sub> dog]] COP barking;
- b.      $\exists y [\lambda n \lambda x [R(x,C) \wedge OU(C)(x) = n](2)(y) \wedge BARKING(y)]$   
 $\Leftrightarrow \exists y [R(y,C) \wedge OU(C)(y) = 2 \wedge BARKING(y)]$

- 515)     [<sub>NP</sub> dog]
- a.     sg: [<sub>NumP</sub> dog]:  $\lambda n \lambda x [R(x,C) \wedge OU(C)(x) = n \wedge n=1];$
- b.     pl: [<sub>NumP</sub> dogs]:  $\lambda n \lambda x [R(x,C) \wedge OU(C)(x) = n$  (possibly:  $n > 1$ )]

Consequently, we have developed our taxonomy of languages with respect to number. The –Num group, whose languages allow bare nouns unspecified for number, splits into two subgroups, i.e. +Cl and –Cl. Languages such as Mandarin Chinese and Vietnamese, namely languages that have classifiers, belong to the +Cl cluster. On the other hand, languages such as creoles belong to the –Cl group. This group further splits into two further clusters on the basis of the animacy requirement which we have observed in languages such as CVC, GBC and Santome (as for the latter, it would be more correct to speak about a +human requirement) as for overt plurality.

Finally, the +Num group is further distinguished into –NumAgr and +NumAgr. Languages which do not require morphological agreement between the noun and the number

word (e.g. Turkish) belong to the –NumAgr group. On the other hand, Romance and Germanic languages belong to the +NumAgr cluster.

## **7.2 Remaining problems**

As for BNs and the nominal system in Kriyol, there are some issues left open. Some of the remaining issues, among others, are a more detailed analysis of specificity issues for indefinite BNs and *un*-NPs, and prosody and phrasing with respect to BN.

## References

- Abney, S. P. 1987. *The English Noun Phrase in its sentential Aspect*. Ph.D. dissertation, MIT.
- Aboh, E.O. 2004. *The Morphosyntax of Complement-Head Sequences: Clause Structure and Word Order Patterns in Kwa*. Oxford, Oxford University Press.
- Aboh, Enoch O. 2009. *D is C: A view from 'genuine' bare noun languages*. University of Amsterdam.
- Aboh, E.O. 2010. The Morphosyntax of the Noun Phrase. In *Topics in Kwa Syntax*, Aboh, Enoch O. & Essegbey, James (eds). Springer Netherlands.
- Alexandre, P. 1972. *An introduction to languages and language in Africa*. London: Heinemann Educational Books.
- Alexandre, N. and Hagemeyer, T. 2007. Bare Nouns and the Nominal Domain in Santome. In *Noun phrases in creole languages: a multi-faceted approach*, Marlyse Baptista & Jacqueline Guéron (eds), 37-59. Amsterdam/ Philadelphia: John Benjamins.
- Allesaib, M. 2012. *Le groupe nominal en Creole Mauricien: etudes syntaxiques (DP syntax in Mauritian)*. Ph.D. dissertation, Universite Paris 8 – Vincennes Saint Denis.
- Amaro, A. 2013. *V Antologia de Poetas Lusófonos*. Leiria: Folheto Edições.
- Arends, J. 1992. Towards a Gradualist Model of Creolization. In *Atlantic Meets Pacific – A Global View Of Pidginization and Creolization (Selected Papers from the Society for Pidgins and Creole Linguistics)*, Francis Byrne and John Holm (eds), 371-380. Amsterdam: John Benjamins.
- Arends, J. 1994. The Socio-historical Background of Creoles. In *Development and Structures of Creole Languages (Essays in Honour of Derek Bickerton)*, 15-24. Amsterdam: John Benjamins.



Arends, J., Bruyn, A. 1994. Gradualist and Developmental Hypothesis. In *Pidgin and Creoles: an Introduction*, Jacques Arends, Pieter Muysken and Norval Smith (eds), 111-120. Amsterdam: John Benjamins.

Arends, J., Kouwenberg, S., and Smith, N. 1994. Theories focusing on the non-European Input. In *Pidgin and Creoles: an Introduction*, Jacques Arends, Pieter Muysken and Norval Smith (eds), 99-110. Amsterdam: John Benjamins.

Arends, J., Muysken, P., and Smith, N. 1994. Conclusions. In *Pidgins and Creoles: an Introduction*, Jacques Arends, Pieter Muysken and Norval Smith (eds), 319-330. Amsterdam: John Benjamins.

Arends, J., Verhagen, J., Lier, E. van, Dikker, S., and Cardoso, H. 2006. On the presence versus absence of morphological marking in four Romance-based Creoles. In *The structure of creole words. Segmental, syllabic and morphological aspects*, Parth Bhatt and Ingo Plag (eds). Niemeyer: Tübingen.

Baker, M. C. 2001. *The Atoms of Language*, New York, Basic Books; trad. it. *Gli atomi del linguaggio: le regole della grammaticale nascoste nella mente*, 2003, Torino, Hoepli.

Baker, P. 1996. Some Developmental Inferences from the Historical Studies of Pidgins and Creoles. In *The Early Stages of Creolization*, Jacques Arends (ed), 1-24. Amsterdam: John Benjamins.

Baker, P. 2000. Theories of Creolization and the Degree and Nature of Restructuring. In *Degrees of Restructuring in Creole Languages*, 41-63. Amsterdam: John Benjamins,

Bakker, P. 1994. Pidgins. In *Pidgins and Creoles: an Introduction*, Jacques Arends, Pieter Myusken and Norval Smith (eds), 25-39. Amsterdam: John Benjamins.

Baptista, M. 2002. *The Syntax of Cape Verdean Creole: The Sotavento Varieties*. Amsterdam/Philadelphia: John Benjamins.

Baptista, M., Mello, H. and Suzuki, M. 2007. The syntax of Cape Verdean Creole and Guinea-Bissau Creole. In *Comparative Creole Syntax*, John Holm and Peter Patrick (eds), 53-82. London: Westminster Creolistics Series, Battlebridge.

- Baptista, M. 2007a. Properties of noun phrases in creole languages: a synthetic comparative exposition. In *Noun Phrases in Creole Languages: A Multi-faceted Approach*, Marlyse Baptista and Jacqueline Guéron (eds), 459-470. Amsterdam/Philadelphia: John Benjamins.
- Baptista, M. 2007b. On the syntax and semantics of DP in Cape Verdean Creole. In *Noun Phrases in Creole Languages: A Multi-faceted Approach*, Marlyse Baptista and Jacqueline Guéron (eds). 61-106. Amsterdam/Philadelphia: John Benjamins.
- Baptista, M. and Guéron, J. 2007a. Noun phrases in creole languages. In *Noun Phrases in Creole Languages: a multi-faceted Approach*, Marlyse Baptista and Jacqueline Guéron (eds). 1-34. Amsterdam/Philadelphia: John Benjamins.
- Baptista, M. and Guéron, J. 2007b. Functional deficiency, ellipsis or innovation in creole languages? In *Noun Phrases in Creole Languages: a multi-faceted Approach*, Marlyse Baptista and Jacqueline Guéron (eds), 471-483. Amsterdam/Philadelphia: John Benjamins.
- Baptista, M. 2013. Cape Verdean Creole of Brava dataset. In *Atlas of Pidgin and Creole Language Structures Online*, Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Barbosa, J. M. 1975. Cape Verde, Guinea-Bissau and São Tomé and Príncipe. The linguistic situation. In *Miscelânea Luso-Africana: colectânea de estudos cologidos*, M. Valkhoff et al. (eds). 133-151. Lisbon: Junta de Investigações Científicas du Ultramar.
- Barbosa, M.J., Conceicao Dias C. and Nafz, M.da. 2009. *Praktisches Lehrbuch Portugiesisch*. Berlin: Langenscheidt.
- Baxter, A.N. 1992. A contribuição das comunidades afro-brasileiras isoladas para o debate sobre a crioulação prévia: um exemplo do estado da Bahia. In *Actas do Colóquio sobre 'Crioulos de Base Lexical Portuguesa'*, Ernesto d'Andrade and Alain Kihm (eds), 7-35. Lisboa: Edições Colibri.
- Benincà, Paola. 1980. Nomi senza articolo. In *Rivista di Grammatica Generativa 5*: 51– 62.

- Bichler, G.A. 2003. *Bejo, Curay und Bin-bim? Die Sprache und Kultur der Wolof im Senegal (mit angeschlossenen Lehrbuch Wolof)*, Europäische Hochschulschriften Band 90, Peterlang Verlagsgruppe, Frankfurt am Main.
- Besten, H. (den), Muysken, P., Smith, N. 1995. Theories focusing on the European Input. In *Pidgins and Creoles: an Introduction*, Jacques Arends, Pieter Muysken and Norval Smith (eds), 87-98. Amsterdam, John Benjamins.
- Bickerton, D. 1981. *Roots of language*. Ann Arbor: Karoma Publishers.
- Bickerton, Derek. 1988. Creole languages and the bioprogram. In *Linguistics: The Cambridge survey. Volume 2: Linguistic theory: Extensions and implications*. Frederick J. Newmeyer, ed., 268-84. Cambridge: Cambridge University Press.
- Bickerton, D. 1990. *Language and Species*. Chicago: University of Chicago.
- Black, M.C., and Gilbert, G.G. 1991. A Reexamination of Bickerton's Phylogenesis Hypothesis. In *Development and Structures of Creole Languages (Essays in Honour of Derek Bickerton)*, 111-122. Amsterdam, John Benjamins.
- Bobyleva, E. 2010. *Variable plural marking and the lexical semantics of (creole) nominal expressions*. Paper presented at FACS II, November 8-9, 2010. Berlin.
- Bobyleva, E. 2011. Definite determiners in two English-based creoles. Specificity or definiteness? In *The Noun Phrase in Romance and Germanic. Structure, variation, and change*. Petra Sleeman and Harry Perridon (eds). Amsterdam/Philadelphia: John Benjamins.
- Bobyleva, E. 2012. *Specificity and topicality effects on plural marking in creoles*. Paper presented at FACS III, November 15-16, 2012, Lisbon.
- Brousseau, Anne-Marie and Lumsden, John S. 1992. Nominal Structure in Fongbe. *The Journal of West African Languages* 22:5-25.
- Bruyn, A. 1995. *Grammaticalization in creoles: the development of determiners and relative clauses in Sranan*. Amsterdam: IFOTT.
- Buis, Pierre. 1990. *Essai sur la langue manjako de la zone de Basserel*. Bissau, Guine-Bissau: Instituto Nacional de Estudos e Pesquisa.

- Capo, Hounkpati B.C. 1988. *Renaissance du Gbe: Réflexions critiques et constructives sur L'EVE, le FON, le GEN, l'AJA, le GUN, etc.* Hamburg: Helmut Buske Verlag.
- Capo, Hounkpati B.C. 1991. *A Comparative Phonology of Gbe*, Publications in African Languages and Linguistics, 14. Berlin/New York: Foris Publications & Garome, Bénin: Labo Gbe (Int).
- Carlson, G.N. 1977. A unified analysis of the English bare plural. *Linguistics and Philosophy* 1 (1977). 413-457. Dordrecht: D. Reidel Publishing Company.
- Castilho, A.T. de, 2007. Prefácio. In *Origens do Português Brasileiro*. Anthony Julius Naro and Maria Marta Pereira Scherre (eds). 11-15. Parábola Editorial, São Paulo.
- Castro, Y. Pessoa de. 2001. *Falares Africanos na Bahia. Um Vocabulário Afro-Brasileiro*. Rio de Janeiro: Academia Brasileira de Letras Topbooks.
- Chafe, W.L. 1976. Givenness, contrastiveness, definiteness, subjects and topics. In *Subject and topic*, Charles N. Li (ed), 27-55. New York: Academic Press.
- Chaudenson, R. 1992. *Des îW les, des hommes, des langues: essais sur la créolisation linguistique et culturelle*. L'Harmattan, Paris. Revised, English edition, *Creolization of Language and Culture*, Routledge: London.
- Cheng, L. Lai-Shen and Sybesma, R. 1999. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30:509–542.
- Cheng, L. Lai.-Shen and Sybesma, R. 2005. Classifiers in four varieties of Chinese. In *The Oxford Handbook of Comparative Syntax*, Guglielmo Cinque and Richard S. Kayne (eds). 259-292. Oxford: Oxford University Press.
- Chierchia, G. 1995. Individual-level predicates as inherent generic. In *The generic book*. Gregory N. Carlson and Francis J. Pelletier (eds), 176-223. Chicago: University of Chicago Press.
- Chierchia, G. 1998. Reference to kinds across languages. *Natural Language Semantics* 6 : 4. 339-400.

- Chomsky, N. 1993. *A Minimalist program for linguistic theory*. In Hale and Keyser (eds), 1-52.
- Chomsky, N. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Christophersen, Paul. 1939. *The Articles: A Study of their Theory and Use in English*. Copenhagen: Einar Munksgaard.
- Claudi, U. 1997. Some thought on the origin of gender marking or Wings were not developed for flying. In *African Linguistics at the Crossroads*. Papers from Kwaluseni. Robert K. Herbert (ed.), 63-74. Cologne: Rüdiger Köppe Verlag.
- Clements, J.C. 1992. Foreigner talk and the origins of Pidgin Portuguese. In *Journal of Pidgin and Creole languages* 7, 75-92.
- Comrie, B., 1976. *Aspect*. Cambridge, Cambridge University Press.
- Condoravdi, A. C. 1997. *Descriptions in context*. New York/London: Garland.
- Cook, V.J. and Newson, M. 2007. *Chomsky's Universal Grammar*. An introduction. Malden/Oxford/Carlton: Blackwell.
- Corbett, G. 1991. *Gender*. Cambridge: Cambridge University Press.
- Corbett, G. 2000. *Number*. Cambridge, UK: Cambridge University Press.
- Couto, H.H. do. 1992. Lançados, grumetes e a origem do crioulo português no nooeste africano. In *Actas do colóquio sobre "Crioulos de base lexical portuguesa*. E. d'Andrade and A.Kihm (eds), 109-122.Lisbon: Colibri.
- Couto, H.H. do. 1996. *Introdução ao estudo das línguas crioulas e pidgins*. Brasília: Editora Universidade de Brasília.
- Couto, H.H. do. 2009. O Português e o crioulo na Guiné-Bissau. In *Português em contacto*, Ana M. Carvalho (ed). Iberoamericana Vervuert: Frankfurt am Main.
- Cyrino, S. and Espinal, M.T. to appear. Bare nominals in Brazilian Portuguese: more on the DP/NP analysis. In *Natural language and linguistic theory*.
- Czypionka, A. 2014. *The interplay of object animacy and verb class in representation building*. PhD dissertation, Humboldt-Universität zu Berlin.

- D'Andrade, E., Gomes Alfredo & Teixeira, Inês. 1992. Observações sobre o sistema accentual do crioulo da Guiné-Bissau (CGB). In *Actas do Colóquio sobre "Crioulos de Base Lexical Portuguesa"*, Ernesto d'Andrade and Alain Kihm (eds), 135-140. Lisbon: Edições Colibri.
- Dayal, V., 1999. *Bare Noun Phrases, Reference to Kinds, and Incorporation*. SALT IX.
- Dayal, V. 2002. Number Marking and (In)Definiteness in Kind Terms. *Linguistics and Philosophy* 27:393-450.
- Dayal, V. 2006. Semantic Variation and Pleonastic Determiners: The Case of the Plural Definite Generic, to appear in *Proceedings of V Asian GLOW*.
- Dayal, V. 2009. Bare Noun Phrases cross-linguistically, to appear in the *Handbook of Semantics*.
- DeGraff, M. 1996. Predicate-Movement, Quantification, Events and Properties. *Linguistics and Philosophy*. MIT, Cambridge MA, USA.
- DeGraff, M. 2001. Morphology in Creole Genesis: Linguistics and Ideology. In *Ken Hale: A Life in Language*, M. Kenstowicz (ed), 53-121. Cambridge MA: MIT Press.
- DeGraff, M. 2005. Morphology and Word Order in 'Creolization' and beyond. In *The Oxford Handbook of Comparative Syntax*, Guglielmo Cinque & Richard Kayne (eds), 293-372. Oxford: Oxford University Press.
- De Wolff, P.P. 1981. Die Afrikanischen Sprachfamilien. I. Niger-Kordofanisch. Das Niger-Kongo (ohne Bantu). In *Die Sprachen Afrikas. Mit zahlreichen Karten und Tabellen*, Bernd Heine, Thilo C- Schadeberg, and Ekkehard Wolff (eds.), 45-76. Hamburg: Helmut Buske Verlag.
- Deprez, V. 2001. On the nature of Bare Nouns in Haitian Creole. In *Current Issues in Linguistics Theory*, D.Cresti. C.Tortora.and T.Satterfield (eds), 47-64. Amsterdam: John Benjamins.
- Deprez, V. 2003. Constraints on the Meaning of Bare Nouns. In *A Romance Perspective on Language Knowledge and Use, Current Issues In Linguistics Theory*

- 238, Raphael Nunes-Cedeno, Luis Lopez, Richard Cameron (eds), 291-311.  
Amsterdam: John Benjamins.
- Deprez, V. 2005. Morphological Number, Semantic Number and bare Nouns. *Lingua*.  
Volume 115, Issue 6:857-883.
- Deprez, V. 2006. On the Conceptual Role of Number. *In New Perspectives on Romance Linguistics*, Nishida, Chiyo and Jean-Pierre Y. Montreuil (eds.), 67-83. Amsterdam: John Benjamins.
- Deprez, V. 2007. Implicit determination and Plural. In *Noun Phrases in Creole Languages: a multi-faceted Approach*, Marlyse Baptista and Jacqueline Guéron (eds). 299-336. Amsterdam/Philadelphia: John Benjamins.
- Depréz, V. 2011. *Specificity and Definiteness in L2 and Creole languages*. Paper presented at the 2011 GRC meeting in Paris.
- Ferreira 2009. Marcadores de plural no português brasileiro e no crioulo cabo-verdiano. In *Português em contacto*, Ana M. Carvalho (ed.). Iberoamericana Vervuert: Frankfurt am Main.
- Derive, Marie-Jo. 1990. *Étude dialectologique de l'aire Manding de Côte-d'Ivoire*. Fascicule 1. Paris: Peeters.
- Derive, Marie-Jo. 1990. *Étude dialectologique de l'aire Manding de Côte-d'Ivoire*. Fascicule 2. Paris: Peeters.
- Dijkhoff, Marta B. 1983. The process of pluralization in Papiamentu. In *Studies in Caribbean Language*, Carrington, Lawrence D., Craig, Dennis & Dandaré, Ramon Todd (eds.), 217-229. St Augustine, Trinidad: Society for Caribbean Linguistics.
- Do Couto, H. H. 1993. The Genesis of Portuguese Creole in Africa. In *Atlantic meets Pacific: a Global View of Pidginization and Creolization*, Francis Byrne and John Holm (eds), 381-389. Amsterdam: John Benjamins.
- Doetjes, J. 2012. Count/mass distinctions across languages. In *Semantics: An International Handbook of Natural Language Meaning*, Part III, C. Maienborn, K. von Stechow, and P. Portner (eds). Berlin: De Gruyter.

- Donati, C. 2002. *Sintassi elementare*. Roma: Carocci.
- Dryer, Matthew S. 2011. Noun-Modifier Order in Africa. In *Geographical Typology and Linguistic Areas. With special reference to Africa*. Osamu Hieda, Christa König, and Hiroshi Nakagawa (eds). 287-311. Amsterdam/Philadelphia: John Benjamins.
- Duarte Inês, Gonçalves Anabela, Miguel Matilde Mendes, Amália Hendrickx, Iris, Oliveira Fátima, Cunha Luís Filipe Silva, Fátima, and Silvano Purificação. 2010. Light verbs features in European Portuguese. In *Proceedings of the Interdisciplinary Workshop on Verbs: The Identification and Representation of Verb Features (Verb 2010)*, Pisa, Italy, Nov.
- Escure, G. 1997. Introduction: Developmental Continua. In *Creole and Dialect Continua*, 1-25. Amsterdam: John Benjamins.
- Escure, G. 1997. Creole Acrolects as Innovations. In *Creole and Dialect Continua*, 57-87. Amsterdam: John Benjamins.
- Fadaïro, D. 2001. *Parlon Fon. Langue et Culture du Bénin*. Paris: L'Harmattan.
- Farkas, D.F. 1994. Specificity and scope. *Langues et Grammaire 1*, L. Nash and G. Tsoulas (eds.).
- Farkas, D.F. 2002. Specificity distinctions. *Journal of semantics* 19 (3), 213-243.
- Farkas, D.F. and de Swart, H. 2007. *The semantics of incorporation: from argument structure to discourse transparency*. University of Chicago Press.
- Fattier, D. 2013. Haitian Creole structure dataset. In *Atlas of Pidgin and Creole Language Structures Online*. Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Ferraz, L.I. 1979. *The Creole of São Tomé*. Johannesburg: Witwatersrand University Press.
- Ferreira, F. 2009. Marcadores de plural no português brasileiro e no crioulo cabo-verdiano. In *Português em Contato*, Ana M. Carvalho (ed), 107-130. Iberoamericana Vervuert: Madrid/Frankfurt am Main.



- Field, F. 2004. Second Language Acquisition in Creole Genesis: the Role of Processability. In *Creoles, Contact and Language Change: Linguistic and Social Implications*, Genevieve Escure and Armin Schwegler (eds), 127-160. Amsterdam: John Benjamins.
- Figueiredo Brauer, F.V. & Brauer, U. 1994. *Praktisches Lehrbuch: Portugiesisch mit Berücksichtigung der brasilianischen Besonderheiten*. Langenscheidt: Berlin.
- Gil, D. 1987. Definiteness, noun phrase, configurationality, and the count-mass distinction. In *The representation of (in)definiteness*, Eric J. Reuland and Alice G.B. ter Meulen (eds.), 254-269. MIT Press: Cambridge, Massachusetts.
- Giorgi, A. & Pianesi, F. 1997. *Tense and Aspect: from Semantics to Morphosyntax*. Oxford: Oxford University Press.
- Gomes, A. and Mendonça, J.G. 1981. *Essai de grammaire créole de Guinée-Bissau*. Bissau: Ministério da Educação.
- Graffi, G. 1994. *Le strutture del linguaggio: Sintassi*. Bologna: Il Mulino.
- Graffi, G. 2008. *Che cos'è la grammatica generativa*. Roma: Carocci.
- Friedlärer, M. 1992. Einführung. In *Lehrbuch des Malinke*. 15-39. Leipzig/Berlin/München: Langenscheidt.
- Greenberg, J.H. 1955. *Studies in African linguistic classification*. New Haven: The Compass Publishing Company.
- Greenberg, J.H. 1966. *Universals of language*. Cambridge, MA: MIT Press.
- Givón, T. 1976. Topic, pronoun, and grammatical agreement. In *Subject and topic*, Charles N. Li (ed.), 151-185. New York: Academic Press.
- Goodman, M. 1987. The Portuguese element in the American Creoles. In *Pidgins and Creole languages: Essays in memory of John E. Reinecke*, Glenn G. Gilbert (ed), 361-405. Honolulu: University of Hawaii Press.
- Greenberg, J.H. 1966. *The languages of Africa*. Bloomington: Indiana University.
- Guéron, J., Lecarme, J. 2004. Introduction. In *The Syntax of Time*, MIT.

- Guéron, J. 2006. Generic sentences and bare plurals. In *Proceedings of the Brussels Conference on Indefinites and Weak Quantifiers*, S. Vogeleer & L. Tasmowski (eds), 219-234. Amsterdam: John Benjamins.
- Guéron, J., Lecarme, J. 2008. Tense and Modality in Nominals. In *Time and Modality, Studies in Natural Language and Linguistic Theory* 75, 195-225. Springer, Dordrecht.
- Guillemin, D. 2011. *The syntax and semantics of a determiner system: a case study of Mauritian Creole*. Amsterdam: John Benjamins.
- Guy, G.R. 1989. On the nature and origins of Popular Brazilian Portuguese. In *Estudios sobre Español de América y Lingüística Afroamericana*, 226-244. Bogotá: Instituto Caro y Cuervo.
- Güldemann, T. 2011. Proto-Bantu and Proto-Niger-Congo. In *Geographical Typology and Linguistic Areas. With special reference to Africa*, Osamu Hieda, Christa König, and Hiroshi Nakagawa (eds), 109-141. Amsterdam/Philadelphia: John Benjamins.
- Tjerk Hagemeyer. 2013. Santome structure dataset. In *Atlas of Pidgin and Creole Language Structures Online*. Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Halle, M., and Marantz, A. 1993. Distributed Morphology and the Pieces of Inflection. In *The View from Building 20*, 112-175. MIT Press.
- Haspelmath, M. 1997. *Indefinite pronouns*. Oxford: Oxford University Press.
- Haspelmath, M. and the APiCS Consortium. 2013. Occurrence of nominal plural markers. *Atlas of Pidgin and Creole Language Structures Online*, Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Heim, I.R. 1982. *The semantics of definite and indefinite noun phrases*. Dissertation: University of Massachusetts.

- Heim, I.R. 1983. File Change Semantics and the Familiarity Theory of Definiteness. In *Meaning, Use, and Interpretation of Language*, R. Bäuerle, Ch. Schwarze, and A.v. Stechow (eds), 164-189. Berlin: de Gruyter.
- Heim, I. 1987. Where does the Definiteness Restriction apply? Evidence from the Definiteness of Variables. In *The representation of (In)definiteness*. Eric J. Reuland and Alice G.B. ter Meulen (eds), 21-42. Cambridge, Mass.: MIT Press.
- Heim, I.R. 1988. *The Semantics of Definite and Indefinite Noun Phrases*. Garland Publishers: New York.
- Higginbotham, J. 1987. Indefiniteness and Predication. In *The representation of (In)definiteness*. Eric J. Reuland and Alice G.B. ter Meulen (eds), 43-70. Cambridge, Mass.: MIT Press.
- Heine, B. and Voßen, R. 1981. Sprachtypologie. In *Die Sprachen Afrikas. Mit zahlreichen Karten und Tabellen*, Bernd Heine, Thilo C- Schadeberg, and Ekkehard Wolff (eds), 407-444. Hamburg: Helmut Buske Verlag.
- Heringer, J. 1969. Indefinite noun phrases and referential opacity. In *Papers from the 5<sup>th</sup> regional meeting of the Chicago Linguistic Society*, hg. v. R.J. Binnick, 89-97. Chicago.
- Holm, J. 1988. Theory and Structure. In *Pidgins and Creoles vol. 1*. Cambridge: Cambridge Language Surveys.
- Holm, J., 1989, Reference Survey. *Pidgins and Creoles vol. 2*. Cambridge: Cambridge Language Surveys.
- Holm, J. 1992. Popular Brazilian Portuguese: a semi-creole. In *Actas do Colóquio sobre "Crioulos de Base Lexical Portuguesa"*, Ernesto d' Andrade and Alain Kihm (eds), 37-66. Edicoes Colibri: Lisboa.
- Holm, J. 2000a. Syntax. In *An introduction to pidgins and creoles*, J. Holm (ed), 171-236. Cambridge: Cambridge University Press.
- Holm, J. 2000b. Conclusions. In *An introduction to pidgins and creoles*. J. Holm (ed), 237-240. Cambridge: Cambridge University Press.

- Holm, J. 2008. Creolization and the fate of inflections. In *Aspects of language contact. New theoretical, methodological and empirical findings with special focus on Romanticisation Processes*, Thomas Stolz, Dik Bakker and Rosa Salas Palomo (eds), 299-324. Berlin/New York: Mouton de Gruyter.
- Hornstein, N. 1990. *As time goes by: Tense and Universal Grammar*. Cambridge: MIT Press.
- Huang, C.-T. James. 1987. Existential sentences in Chinese and (in)definiteness. In *The representation of (in)definiteness*, Eric Reuland and Alice ter Meulen (eds), 226–253. Cambridge: MIT Press.
- Huang, C.-T. James, Li, Y.-H. Audrey, and Li, Yefei. 2009a. The *ba* construction. In *The syntax of Chinese*, 153-196. Cambridge: Cambridge University Press.
- Huang, C.-T. James, Li, Y.-H. Audrey, and Li, Yefei. 2009b. Nominal expressions. In *The syntax of Chinese*, 283-328. Cambridge: Cambridge University Press.
- Hull, A. 1993. The Transmission of Creole Languages. In *Atlantic Meets Pacific – A Global View Of Pidginization and Creolization (Selected Papers from the Society for Pidgins and Creole Linguistics)*, Francis Byrne and John Holm (eds), 391-397. Amsterdam: John Benjamins.
- Huttar, G. 1991. Ndjuka Organization of Experience: African or Universal? In *Development and Structures of Creole Languages (Essays in Honour of Derek Bickerton)*, Francis Byrna and Thom Huebner (eds), 101-111. Amsterdam: John Benjamins.
- Incopté, E. 2012. *Insana Rebeldia*. Coimbra: Temas originais.
- Intumbo, I. 2007. *Estudo comparativo da morfossintaxe do crioulo guineense, do balanta e do português*. M.A. thesis. Universidade de Coimbra.
- Intumbo, I. 2008. Balanta, Guiné-Bissau Creole Portuguese and Portuguese. A comparison of the noun phrase. In *Roots of creole structures. Weighing the contribution of substrates and superstrates*. Susanne Michaelis (ed), 263-278. Amsterdam: John Benjamins.
- Intumbo, I., Inverno, L. and Holm, J. 2013. Guinea-Bissau Kriyol structure dataset.

In *Atlas of Pidgin and Creole Language Structures Online*, Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds).

Leipzig: Max Planck Institute for Evolutionary Anthropology.

Ionin, T. 2006. *This is definitely specific: specificity and definiteness in article systems*. *Natural Language Semantics*, 14, 175-234.

Jacobs, B. 2009. The Upper Guinea origins of Papiamentu. Linguistic and historical evidence. *Diachronica* 26:3. 319-379.

Jacobs, B. 2010. Upper Guinea Creole. Evidence in favor of a Santiago birth. *Journal of Pidgin and Creole Languages* 25:2 (2010), 289–343.

Kamp, H. 1981. A theory of truth and semantic representation. In *Formal Methods in the Study of Language*, J.A.G. Groenendijk, T.M.V. Janssen, and M.B.J. Stokhof (eds.), 277-322. Mathematical Centre Tracts 135, Amsterdam.

Karttunen, L. 1971. *The logic of English predicate complement constructions*.

Publications of the Indiana University Linguistics Club, Bloomington.

Karttunen, L. 1976. Discourse referents. In *Syntax and semantics 7: Notes from the linguistic underground*. 363-385. J.D. McCawley (ed.), Academic Press, New York.

Kaye, A.S. and Tosco, M. 2001. *Pidgin and Creole Languages: a basic Introduction*. Munich: Lincom Europa, Lincom Textbook in Linguistics.

Kester, E. & Schmitt, C. 2007. Papiamentu and Brazilian Portuguese: a comparative study of bare nominal. In *Noun Phrases in Creole Languages: a multi-faceted Approach*, Marlyse Baptista and Jacqueline Guéron (eds), 107-143.

Amsterdam/Philadelphia: John Benjamins.

Kihm, A. 1992. Funny objects in Kriyol. In *Actas do Colóquio sobre “Crioulos de Base Lexical Portuguesa”*, Ernesto d’Andrade and Alain Kihm (eds), 123-134.

Lisbon: Edições Colibri.

Kihm, A. 1993. What is it that you said? A study of obligatory focalization in two creoles and beyond. In *Focus and the grammatical relation in creole languages*, Francis Byrne and Donald Winford (eds), 141-162. Amsterdam: John Benjamins.

- Kihm, A. and Lefebvre, C. 1993. *Aspects de la grammaire du Fongbé. Études de phonologie, de syntaxe et de sémantique*. Paris: Peeters Press.
- Kihm, A. 1994. *Kriyol Syntax: the Portuguese-based Creole Language of Guinea-Bissau*. Amsterdam: John Benjamins.
- Kihm, A. 1999. Focus in Wolof: a study of what morphology may do to syntax. In *The Grammar of Focus*, G. Rebuschi & L. Tuller (eds), 245-73. Amsterdam: John Benjamins.
- Kihm, A. 2000. Are Creole Languages Perfect Languages? In *Language Change and Language Contact in Pidgins and Creoles*, John McWhorter (ed), 163-200. Amsterdam: John Benjamins.
- Kihm, A. 2002. *What's in a noun: Noun Classes, Gender, and Nounness*. Ms., Université Paris 7.
- Kihm, A. *Interpreting fully bare count nouns in creole languages*, (Laboratoire de Linguistique formelle, CNRS - Université Paris 7).
- Kihm, A. 2007a. On the interpretation of bare noun phrases in Guinea-Bissau Portuguese Creole (Kriyol). In *Noun Phrases in Creole Languages: a multi-faceted Approach*, Marlyse Baptista and Jacqueline Guéron (eds), 145-169. Amsterdam/Philadelphia: John Benjamins.
- Kihm, A., 2007b. The two faces of creole grammar and their implications for the origin of complex language. In R. Eckardt, G. Jäger, & T. Veenstra (eds.), *Variation, selection, development: Probing the evolutionary model of language change*, 253-305. Berlin: Mouton de Gruyter.
- Kihm, A. 2008. Creolization and Markedness Theory. In *The Handbook of Creole Linguistics*, S. Kouwenberg and J. Singler (eds). Oxford: Basil Blackwell.
- Kihm, A. 2010. *Comparing the inflectional morphologies of Portuguese-related Creoles: a realizational and SLA-based account*. Papers presented at FACS II, Berlin, 8-9 November 2010.

- Kihm, A. 2011. Substrate influence in Kriyol: Guinea-Bissau and Casamance Portuguese-related Creole. In *Creoles, their Substrates, and Language Typology*. Claire Lefebvre (ed.). 81-103. Amsterdam: John Benjamins.
- Kihm, A. & Rougé, J.-L. To appear. *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. In *Journal of Pidgin and Creole Languages*.
- Kirby, P.J. 2006. *Vietnamese and the Structure of Noun Phrase*, Ms. University of Chicago.
- Klein, W. 1994. *Time in language*. London: Routledge.
- Kouwenberg, S. & Murray, E. 1994. Papiamentu. *Languages of the world/Materials* 83 München: Lincom Europa.
- Kouwenberg, S. & Ramos-Michel, A. 2007. Papiamentu (Creole Spanish/Portuguese). In *Comparative Creole Syntax*, John A. Holm and Peter L. Patrick (eds), 307-332. Westminster: Battlebridge.
- Kouwenberg, S. 2013. Papiamentu structure dataset. *Atlas of Pidgin and Creole Language Structures Online*, Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Kramer, J. 2005. Die Entstehung des Papiamentu aus spanisch-portugiesischer Konvergenz in niederländischem Munde. In *Portugiesische Kreolsprachen: Entstehung, Entwicklung, Ausbau und Verwendung*. Annette Endruschat and Axel Schönberger (eds.). Domus Editoria Europaea: Frankfurt am Main.
- Krasnoukhova, O. 2012. *The Noun Phrase in the Languages of South America*. PhD dissertation. Radboud University: Nijmegen.
- Kratzer, A. 1995. Stage-level and individual-level predicates. In *The generic book*, Gregory N. Carlson and Francis J. Pelletier (eds), 125-175. Chicago: University of Chicago Press.

- Krifka, M. 1989. Nominal reference, temporal constitution and quantification in event semantics. In *Semantics and contextual expression*, Renate Bartsch, Johann van Benthem, and Peter van EmdeBoas (eds), 75-115. Stanford, CA: CSLI Publications.
- Krifka, M., Pelletier, F.J., Carlson G.N., Meulen, A. ter, Link G., and Chierchia G. 1995a. 'Genericity: an introduction'. In *The generic book*, Gregory N. Carlson and Francis J. Pelletier (eds), 1-124. Chicago: University of Chicago Press.
- Krifka, M. 1995b. Focus and the interpretation of generic sentences. In *The generic book*, Gregory N. Carlson and Francis J. Pelletier (eds), 238-264. Chicago: Chicago Press.
- Krifka, M. 1995c. Common Nouns: a contrastive Analysis of Chinese and English. In *The generic book*, Gregory N. Carlson and Francis J. Pelletier (eds), 398-411. Chicago: Chicago Press.
- Krifka, Manfred. 2004. Bare NPs: kind-referring, indefinites, both, or neither? In *Proceedings of Semantics and Linguistic Theory (SALT) 13*, Rob Young & Yuping Zhou (eds). Cornell U: CLC Publication.
- Krifka, M. 2007. Basic notions of information structure. In *Interdisciplinary Studies on Information Structure 6*, Caroline Fery & Manfred Krifka (eds), 13-56. Potsdam: Universitätsverlag.
- Lanciani, G., and Tavani, G. 1993. *Grammatica Portoghese*. Edizioni Universitarie di Lettere, Economia e Diritto: Milano.
- Lang, G. 2000. Chaos and Creoles: towards a New Paradigm? In *Creolization and Contact in Pidgins and Creoles*, John McWhorter (ed), 443-458. Amsterdam: John Benjamins.
- Lang, J. 2013. Cape Verdean Creole of Santiago structure dataset. In *Atlas of Pidgin and Creole Language Structures Online*, Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Lecarme, J. 2004. Tense in Nominals. In *The Syntax of Time*, 441-475. MIT.



- Lefebvre, C. 1998. The Problem of Creole Genesis and Linguistic Theory. In *Creole Genesis and the Acquisition of Grammar: the Case of Haitian Creole*. Cambridge: Cambridge University Press.
- Lefebvre, C. and Brousseau, A.-M. 2002. *A grammar of Fongbe*. Berlin: Mouton de Gruyter.
- Lefebvre, C. 2011. The problem of the typological classification of creoles. In *Creoles, their substrates and language typology*, Claire Lefebvre (ed), 3-33. Amsterdam/Philadelphia: John Benjamins.
- Lewis, J. 2000. "The noun. In *Turkish Grammar*, 23-49. Oxford: Oxford University Press.
- Li, Charles, and Sandra Thompson. 1975. The semantic function of word order: A case study in Mandarin. In *Word order and word order change*, Charles N. Li (ed), 163–186. Austin: University of Texas Press.
- Li, Charles, and Sandra Thompson. 1989a. Introduction. In *Mandarin Chinese: A functional reference grammar*, 1-9. University of California Press and Los Angeles: University of California Press.
- Li, Charles, and Sandra Thompson. 1989b. Typological description. In *Mandarin Chinese: A functional reference grammar*, 10-27. University of California Press and Los Angeles: University of California Press.
- Li, Charles, and Sandra Thompson. 1989c. Simple declarative sentences. In *Mandarin Chinese: A functional reference grammar*. University of California Press and Los Angeles: University of California Press.
- Li, Charles, and Sandra Thompson. 1989d. The ba construction. In *Mandarin Chinese: A functional reference grammar*. University of California Press and Los Angeles: University of California Press.
- Link, G. 1983. The logical analysis of plurals and mass terms: a lattice-theoretic approach. In *Formal semantics – the essential readings*, P. Portner and B.H. Partee (eds), 127-147. Blackwell.

- Link, 1987. Generalized quantifiers and plurals. In *Generalized quantifiers, linguistic and logic approaches*, P. Gärdenfors (ed.). D. Reidel Publishing Company.
- Lipski, J.M. 2009. *Os primeiros contatos afro-portugueses: implicações para a expansão da língua*. Iberoamericana Vervuert.
- Longobardi, Giuseppe. 1994. Reference to Kinds and proper names: a theory of N-movement in syntax and Logical form. *Linguistic Inquiry* 25: 609-665.
- Longobardi, G. 1996. The syntax of N-raising: a minimalistic theory. *OTS Working Papers*, 1-55. Utrecht.
- Longobardi, G. 2000. “Postverbal” subjects and the mapping hypothesis. *Linguistic Inquiry* 31: 4, 691-702.
- Lopes, R. E. V. 2005. Bare nouns and DP number agreement in the acquisition of Brazilian Portuguese. In *Selected Proceedings of the 9th Hispanic Linguistics Symposium 1*: 252-262. State College, PA. Somerville, MA: Cascadilla Press.
- Lucchesi, Dante. 1993. The article systems of Cape Verde and Sao Tomé Creole Portuguese: General principles and specific factors. *Journal of Pidgin and Creole Languages* 8, 81-108.
- Lucchesi, D. 1998. A constituição histórica do português brasileiro como um processo bipolarizador: tendências atuais de mudança nas normas culta e popular. In “*Substandard*” e mudança no português do Brasil, Sybille Große and Klaus Zimmermann (eds), 73-99. Frankfurt am Main: TFM.
- Lyons, C. 1999. *Definiteness*. Cambridge University Press.
- Maduro, A.J. 1966. *Procedencia de palabranan Papiamenti i otro anotacionnan*. Curaçao, Netherlands Antilles: privately printed.
- Manzini, M.R. and Savoia, L.M. 2008. *Worknotes on Romance Morphosyntax: Appunti di Morfosintassi Romanza*. Alessandria: Edizioni dell’Orso.
- McWhorter, J. 2001. The world’s simplest grammars are creole grammars. *Linguistic Typology* 5:2-3, 125-166.
- McWhorter, J. 2005. *Defining Creole*. Oxford: Oxford University Press.

- Märzhäuser, C. 2013. Coordinated bare nouns in French, Spanish and European Portuguese. In *New perspectives on bare noun phrases in Romance and beyond*, Johannes Kabatek and Albert Wall (eds), 283-300. John Benjamins: Amsterdam/Philadelphia.
- Matthews, P.H. 2007. Oxford. *Coincise dictionary of Linguistics*. Oxford: Oxford University Press.
- Matthews, S. 1993. Creole Aspect and Morphological Typology. In *Atlantic Meets Pacific – A Global View Of Pidginization and Creolization (Selected Papers from the Society for Pidgins and Creole Linguistics)*, Francis Byrne and John Holm (eds), 233-241. Amsterdam: John Benjamins.
- Matushansky, O. 2005. Call me Ishmael. In E. Maier, C. Bary, and J. Huitink (eds), *Proceedings of Sinn und Bedeutung 9*. Nijmegen: NCS.
- Maurer, P. 1997. Tense-Aspect-Mood in Principense. In *The Structure and Status of Pidgins and Creoles*, Arthur K. Spears and Donald Winford (eds), 415-435. Amsterdam: John Benjamins.
- Maurer, P. and the APiCS Consortium. 2013. Gender agreement of adnominal adjectives. *Atlas of Pidgin and Creole Language Structures Online*, Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Mazzoli, M. 2013. *Copulas in Nigerian Pidgin*. Ph.D. dissertation. University of Bologna, Italy.
- Märzhäuser, C. 2013. Coordinated bare nouns in French, Spanish and European Portuguese. In *New perspectives on bare noun phrases in Romance and beyond*, Johannes Kabatek and Albert Wall (eds), 283-300. Amsterdam: John Benjamins.
- Mello, M.A.C. da Rocha de. 2007. *A questão da produtividade morfológica no Guineense*. PhD-thesis. University of Brasilia.
- Michaelis, S.M., Maurer, P., Haspelmath, M. & Huber, M. (eds). 2013. *The atlas of pidgin and creole language structures*. Oxford: Oxford University Press.

- Michaelis, S.M., Maurer, P., Haspelmath, M. & Huber, M. (eds). 2013. *The survey of pidgin and creole languages, Volumes 1-3*. Oxford: Oxford University Press.
- Mikkelsen, L. 2005. Partee's theory of Noun Phrase interpretation. In *Copular Clauses: Specification, Predication and Equation*, 53-54. Amsterdam: John Benjamins.
- Milsark, G. 1979. *Existential sentences in English*. New York NY: Garland.
- Montague, R. 1974. The proper treatment of quantification in ordinary English. In *Formal Philosophy: Selected Papers of Richard Montague*, R.H. Thomason (ed). New Haven CT: Yale University Press.
- Mufwene, S. 1991. Pidgins, Creoles Typology and Markedness. In *Development and Structures of Creole Languages (Essays in Honour of Derek Bickerton)*, Francis Byrne and Thom Huebner (eds), 123-144. Amsterdam: John Benjamins.
- Mufwene, S. 1996. The Founder Principle in creole genesis. *Diachronica* 13:83-134.
- Myers-Scotton, Carol. 2002. *Contact Linguistics: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
- Mühlhäusler, P. 1986. *Pidgin and Creole linguistics*. Oxford: Blackwell Publishing.
- Müller, A. and Oliveira, F. 2002. Bare Nominals and Number in Brazilian and European Portuguese. Paper presented at the 3<sup>o</sup> Colóquio Português Europeu e Português Brasileiro – Unidade e Diversidade na Passagem do Milénio, held in 2002 in Lisbon, Portugal.
- Munn, A., & Schmitt, C. 1999. *Bare Nouns and the Morphosyntax of Number*. Paper presented at LSRL 29, University of Michigan, April 1999.
- Muysken, P., Voort, H. van den. 1991. The Binding Theory and Creolization: Evidence from 18th Century Negerhollands Reflexives. In *Development and Structures of Creole Languages (Essays in Honour of Derek Bickerton)*, Francis Byrne and Thom Huebner (eds), 145-158. Amsterdam: John Benjamins.
- Muysken, P., and Smith, N. 1994. Reflexives. In *Development and Structures of Creole Languages (Essays in Honour of Derek Bickerton)*, Francis Byrne and Thom Huebner (eds), 271-288. Amsterdam: John Benjamins.

- Muysken, P., Smith, N. 1994. The Study of Pidgin and Creole Languages. In *Pidgin and Creoles: an Introduction*, Jacques Arends, Pieter Muysken e Norval Smith (eds), 3-14. Amsterdam: John Benjamins.
- Muysken, P., Veenstra, T. 1995. Universalist Approaches. In *Pidgin and Creoles: an Introduction*, Jacques Arends, Pieter Muysken e Norval Smith (eds), 121-134. Amsterdam: John Benjamins.
- Muysken, P. 2000. Wolof-French and Fongbe-French code-mixing. In *Bilingual Speech: a typology of code-mixing*, 87-90. Cambridge: Cambridge University Press.
- Naro, A. 1978. A study on the origins of pidginization. *Language* 54: 314–347.
- Naro, A.J. and Scherre, M.M.P. 2007. Sobre as origens do Português popular do Brasil. In *Origens do Português Brasileiro*, Anthony Julius Naro and Maria Marta Pereira Scherre (eds), 25-48. Parábola Editorial: São Paulo.
- Naro, A.J. and Scherre, M.M.P. 2007. Concordância variável em Português: a situação no Brasil e em Portugal. In *Origens do português brasileiro*, Anthony Julius Naro and Maria Marta Pereira Scherre, 49-69. São Paulo: Parábola Editorial.
- Ndiaye, Moussa D. 2004. *Éléments de morphologie du wolof. Méthodes d'analyse en linguistique*. Munich: LINCOM Studies in African Linguistics.
- Nguyen, C. 2012. *The independence of specificity types in Vietnamese*. Paper presented at LSA 86th Annual Meeting, 2012 January 7.
- Nguyên, T.H. 2004. *The Structure of the Vietnamese Noun Phrase*. PhD-dissertation. Boston University.
- Nguyên, T.H. 2008. *The structure of the Vietnamese noun phrase*. VDM Verlag Dr. Müller: Saarbrücken.
- Castro, Y. Pessoa de. 2001. *Falares Africanos na Bahia. Um Vocabulário Afro-Brasileiro*. Academia Brasileira de Letras: Topbooks.
- Oliveira, F. and Silva, F. 2008. Indefinites in generic contexts in European Portuguese. Paper presented at *Actas del VIII Congreso de Linguística General*, Madrid.

- Partee, B.H. and Rooth, M. 1983. Generalized conjunction and type ambiguity. In *Meaning, use, and interpretation*, R. Bäuerle, C. Schwarze, and A. von Stechow (eds), 361-383. Walter de Gruyter: Berlin.
- Partee, B.H. 1986. Noun phrase interpretation and type-shifting principles. In *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*, Jeroen Groenendijk, Dick de Jongh, and Martin Stokhof (eds), 115-143. Dordrecht: Foris.
- Pinto Bull, B. 1989. *O crioulo de Guiné-Bissau: filosofia e sabedoria*. Lisboa/Bissau: Instituto de Cultura e Língua portuguesa (Portugal)/Instituto nacional de estudos e pesquisas (Guiné-Bissau).
- Postal, P. 1969. Anaphoric islands. In *Papers from the Fifth regional meeting of the Chicago Linguistic Society*, Binnick et al. CLS.
- Prince, E.F. 1981. On the interfacing of indefinite-*this* NPs. In *Elements of Discourse Understanding*, A.K. Joshi, B.L. Webber & I.A. Sag (eds), 231-250. Cambridge: CUP.
- Quine, W.V.O. 1960. *Word and object*. MIT Press.
- Reichenbach, H. 1947. *Elements of symbolic logic*. New York: Macmillan.
- Reuland, E.J. and ter Meulen, A.G.B. 1987. Introduction. In *The representation of (In)definiteness*, Eric J. Reuland and Alice G.B. ter Meulen (eds), 1-20. Cambridge: MIT Press.
- Rizzi, L. 1997. The fine structure of the left periphery. In *Elements of Grammar: Handbook in Generative Syntax*, L. Haegeman (ed), 281-337. Dordrecht: Kluwer.
- Rooij, V. 1994. Variation. In *Pidgins and Creoles: an Introduction*, Jacques Arends, Pieter Muysken and Norval Smith (ed), 53-64. Amsterdam: John Benjamins.
- Rougé, J.-L. 1986. *Uma hipótese sobre a formação do crioulo da Guiné-Bissau e da Casamansa*. Soronda 2:28-49.
- Rooth, M. 1985. *Association with Focus*. PhD thesis, University of Massachusetts, Amherst, GLSA, Dept. of Linguistics, South College, UMASS, Amherst MA 01003.

- Russell, B. 1905. On denoting. *Mind* 14: 479–493.
- Scherre, M.M.P. 2007. Introdução. In *Origens do Português Brasileiro*, Anthony Julius Naro and Maria Marta Pereira Scherre (eds), 17-23. Parábola Editorial: São Paulo.
- Schmitt, C. 1996. *Aspect and the Syntax of Noun Phrases*. Doctoral dissertation, University of Maryland, College Park.
- Schmitt, C. & Munn, A. 1999. *Against the Nominal Mapping Parameter: Bare nouns in Brazilian Portuguese*. Proceedings of NELS 29.
- Schmitt, C. & Munn, A. 2000. *Bare Nominals, Morphosyntax, and the Nominal Mapping Parameter*. Ms. Michigan State University.
- Schmitt, C. and A. Munn. 2003. The syntax and semantics of bare arguments in Brazilian Portuguese. *Linguistic Variation Yearbook* 2, 185-216.
- Schuchardt, H. 1888. Beiträge zur Kenntnis des kreolischen Romanisch II: zum Negerportugiesichen Senegambiens. In *Zeitschrift für romanische Philologie* 12, 301-322.
- Schwarz, F. 2009. *Two Types of Definites in Natural Language*. PhD dissertation, University of Massachusetts Amherst, Amherst, MA.
- Seuren, P. and Wekker, H. 1986. Semantic transparency as a factor in creole genesis. In *Substrata versus universals in creole genesis*, Pieter Muysken and Norval Smith (eds), 57-70. Amsterdam: John Benjamins.
- Siegel, J. 1997. Mixing, Leveling, and Pidgin/Creole Development. In *The Structure and Status of Pidgins and Creoles*, Arthur K. Spears and Donald Winford (eds), 111-149. Amsterdam: John Benjamins.
- Singler, J.V. 1990. Introduction: Pidgins and Creoles and Tense-Mood-Aspect. In *Pidgin and Creole Tense-Mood-Aspect Systems*, J.V. Singler (ed). Amsterdam: John Benjamins.
- Slomanson, P.A. 1993. Hesseling and Van Ginneken on Language Contact, Variation and Creolization. In *Atlantic Meets Pacific – A Global View Of Pidginization and*

- Creolization (Selected Papers from the Society for Pidgins and Creole Linguistics)*, Francis Byrne and John Holm (eds), 419-428. Amsterdam: John Benjamins.
- Smith, P. 2003. *An introduction to formal logic*. Cambridge: Cambridge University Press.
- Storch, A. 1995. Einführung und Grundlagen. In *Die Anlautpermutation in den Westatlantischen Sprachen*. Frankfurter Afrikanistische Blätter, Sondernummer 2.
- Stassen, Leon. 1997. The nominal strategy. In *Intransitive Predication*, 63-120. Oxford: Oxford University Press.
- Stowell, T. 1989. Subjects, specifiers, and X-bar theory. In *Alternative conceptions of phrase structure*, M. Baltin & A.S. Kroch (eds), 232-262. Chicago IL: Chicago University Press.
- Strawson, P.F. 1950. On referring. *Mind* 59:320-344.
- Swolkien, D. 2013. Cape Verdean Creole of São Vicente structure dataset. In *Atlas of Pidgin and Creole Language Structures Online*, Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Szabolcsi, A. 1994. The noun phrase. In *The syntactic structure of Hungarian. Syntax and semantics 27*, Ferenc Kiefer & Katalin Kiss (eds), 179–274. San Diego: Academic Press.
- Szabolcsi, A. 2010. *Quantification*. Cambridge: Cambridge University Press.
- Taylor, D. 1961. New languages for old in the West Indies. *Comparative Studies in Society and History* 3:277-288.
- Thomason, S. G. 1997. A typology of Contact Languages. In *The Structure and Status of Pidgins and Creoles*, Arthur K. Spears & Donald Winford, 71-88. Amsterdam: John Benjamins.
- Thompson, R.W. 1961. A note on some possible affinities between the creole dialects of the Old World and those of the New. In *Creole Language Studies 2*, R.B. Le Page (ed), 107-113. London: Macmillan.



- Trinh, T. 2011. Nominal reference in two classifier languages. In *Proceedings of Sinn & Bedeutung 15*, I. Reich et al. (eds), 1–16. Universaar – Saarland University Press: Saarbrücken, Germany.
- Truppi, C. 2009. *Alcuni aspetti della Sintassi del Creolo di Guinea-Bissau*. M.A. thesis, University of Florence, Italy.
- Truppi, C. 2013. *Bare Nouns in Kriyol and the role of specificity*. Paper presented at the ACBLPE & SPCL Joint Meeting, Universidade de Lisboa, Portugal. 19-21 June 2013.
- Truppi, C. 2014. *The null/overt copula in creoles: the case of Guinea-Bissau Creole*. Paper presented at the Workshop on Copulas, University of Bologna, Italy.
- Veenstra, T. 2008. Creole genesis: the impact of the Language Bioprogramm Hypothesis. In *Handbook of Pidgin and Creole Studies*, S. Kouwenberg & J.V. Singler (eds). Malden, MA: Blackwell.
- Veiga, M. 1987. *Odju d'Agu*. Praia: Instituto Caboverdiano do Livro.
- Velupillai, V. 2015. *Pidgins, Creoles and Mixed Languages. An introduction*. Amsterdam/Philadelphia: John Benjamins.
- Vendler, Z. 1971. Singular terms. In *Semantics: An interdisciplinary reader in philosophy, linguistics and psychology*, D.D. Steinberg & L.A. Jakobovits (eds), 115-133. Cambridge: CUP.
- Vergnaud, J.-R. and M. L. Zubizarreta. 1992. The definite determiner in French and English. *Linguistic Inquiry* 23: 595-652.
- Vigh, H. 2006. *Navigating terrains of war: Youth and soldiering in Guinea-Bissau*. New York NY: Berghahn, Books.
- Williamson, K. and Blench, R. 2000. Niger-Congo. In *African languages. An introduction*. Berns Heine and Derek Nurse (eds), 11-42. Cambridge: Cambridge University Press.

Winford, D. 1997. Introduction: on the Structure and Status of Pidgins and Creoles. In *The Structure and Status of Pidgins and Creoles*, Arthur K. Spears and Donald Winford (eds), 1-31. Amsterdam: John Benjamins.

Winford, D., and Plag, I. 2013. Sranan structure dataset. In *Atlas of Pidgin and Creole Language Structures Online*, Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds). Leipzig: Max Planck Institute for Evolutionary Anthropology.

Zampirelli, R. 2000. *Layers in the determiner phrase*. New York NY: Garland.