LILIANA CRISTINA CORAGEM INVERNO

# Angola's Transition to Vernacular Portuguese

Faculdade de Letras
Universidade de Coimbra
2005

LILIANA CRISTINA CORAGEM INVERNO

# Angola's Transition to Vernacular Portuguese

Dissertação de Mestrado na área científica de Linguística Descritiva, área de especialização em Línguas em Contacto: Pidgins, Crioulos e Semi-Crioulos apresentada à Faculdade de Letras da Universidade de Coimbra, sob a orientação do Professor Doutor John A. Holm.

Faculdade de Letras
Universidade de Coimbra
2005

*For my family*

# Contents


ACKNOWLEDGEMENTS *page* i

LIST OF TABLES iii

LIST OF MAPS AND FIGURES iv

ABBREVIATIONS AND SYMBOLS v

ABSTRACT vii

**CHAPTER 1: Introduction** 1

1.0. *Introduction* 1

1.1. *Object of study and aims of the study* 7

1.2. *Structure of the thesis* 8

1.3. *The data* 9

**CHAPTER 2: Overview of the Literature** 11

2.0. *Introduction* 11

2.1. *The study of Angolan Vernacular Portuguese* 11

2.1.1. Reference works in linguistics 12

2.1.2. General works about varieties of Portuguese around the world 14

2.1.3. AVP in Portuguese handbooks and grammars 22

2.1.4. Works on Angola's Bantu languages 29

2.1.5. Specific works on Angolan Vernacular Portuguese 33

2.2. *Contact Linguistics: an overview* 48

2.2.1. Key concepts in pidgin and creole linguistics: definitions 48

2.2.2. Contact linguistics: the theoretical foundations 51

2.2.3. Continuum of contact languages and partial restructuring 56

2.3. *Summary and conclusions* 58

**CHAPTER 3: AVP - The Sociolinguistic Setting of its Development** 60

3. 0. *Introduction* 60

3.1. *Portuguese overseas expansion until the discovery of the Congo* 61

3.2. *The early contacts (1482-1565)* 62

3.2.1. The Congo Kingdom 62

3.2.2. The Ndongo Kingdom 66

3.3. *The founding of Luanda: the early days of the colony of Angola (1575-1640)* 67

3.4. *The Dutch capture of Angola and Brazilian rule of the colony (1641-1700)* 71

3.4. *Angola in the 18th century* 73
3.4.1. The development of Luanda 73
3.4.2. The exploration of the south 75
3.5. *The 19th century and the start of Portugal's African Empire: the case of Angola* 77
3.6. *The 20th century* 84
3.6.1. The pacification wars, the Republic and the early years of the *Estado Novo* (1900 – 1930) 84
3.6.2. From the 1950s to 1975 89
3.7. *Summary and Conclusions* 94

**CHAPTER 4: Contact-induced Features in the AVP Noun Phrase** 96

4.0. *Introduction* 96
4.1. *The noun phrase in European Portuguese* 97
4.1.1. Internal structure and word order 98
4.1.2. Number and gender marking and agreement 100
4.1.3. Possession 103
4.1.4. Case and word order of person markers 104
4.2. *The noun phrase in Kikongo, Kimbundu, Umbundu and Cokwe* 108
4.2.1. Internal structure and word order 114
4.2.2. Number marking strategies 116
4.2.3. Gender marking strategies 121
4.2.4 Case and word order of person markers 125
4.2.5. Possession 129
4.3. *The noun phrase in Angolan Vernacular Portuguese* 130
4.3.1. Internal structure and word order 130
4.3.2. Number marking and agreement 130
4.3.3. Gender marking and agreement 135
4.3.4. Possession 139
4.3.5. Case and word order of person markers 141
4.4. *Summary and conclusions* 143

**CHAPTER 5: The Development of AVP – Conclusions** 145

5.1. *Introduction* 145
5.2. *The development of AVP: a tentative hypothesis* 145
5.3. *Further research* 151

**REFERENCES** 152

# Acknowledgements

*Kasipu okuvala?*
What is it that cannot be counted?

*Enyofi.*
The stars[1]

All transitional stages in one's life are difficult and complex, but the transition from being a student to becoming a salaried worker is a particularly difficult one, especially when one has no wish to give up studying. The choice is usually between pursuing one's dream and securing one's subsistence. I chose both and, like so many other graduate students, had to cope with the consequences of that choice. I have lost track of the moments I thought I would not be able to start this thesis, let alone finish it. It was the friendship, support and patience of my family, friends, colleagues and teachers that kept me sane enough to proceed with the research that made this thesis possible. I am particularly indebted to the following people and institutions:

Professor John Holm, for having accepted to be my supervisor, for all his criticisms, suggestions and help during the writing of this thesis, especially in the final stage, but above all for having believed from the very beginning that I would be able to write it in spite of my own disbelief. It is his unflagging support of the work of young linguists together with the high standards of his own research that has put him among the leading figures in contemporary contact linguistics;

Professors Clara Keating and Isabel Pereira and all my fellow masters' students for their unconditional support, especially Dominika Swolkien and Olga Solovova, for their unflagging friendship and support as well as their insightful comments on my work; Michael Pye for letting me use his personal library, for bringing literature to me that was inaccessible in Portugal, for sharing his knowledge of the history of the diplomatic relations between the African and European rulers in past centuries, and above all for having always been on my side;

Professor Jorge Morais Barbosa and Ana Cristina Macário Lopes, for being an inspiration to me during my undergraduate studies, but also for supporting me in obtaining the grant I was awarded by the Fundação para a Ciência e Tecnologia; Alan Baxter, Clancy Clements, Jürgen Lang, Mauro Fernandez, Nélia Alexandre, Perpétua Gonçalves, Philippe Maurer, Tjerk Hagemeijer and Tom Gueldemann for their valuable insights on my work at ACBLPE meetings since 2001;

[1] Ambó riddle in Oliveira (2001:100).

Antónia, for giving up her Sunday afternoons for several months to teach me Kimbundu, and Domingos, for sharing his experience growing up a native speaker of Kikongo in a country where only Portuguese was considered a "civilized" language; Professor José Manuel Silva, for arranging the cooperation agreement that enabled me to go to Angola in July and August 2003 to collect the data in this thesis on the vernacular Portuguese spoken in Dundo; my colleagues and students in Dundo, for their hospitality, for sharing precious linguistic information on the use of their mother tongues (Kikongo, Cokwe and Umbundu) and for being kind enough to share their first-hand experiences of the sociolinguistic setting in Angola since the 1960s; Professor Américo Oliveira for showing me around Luanda and the lovely *Barra do Cuanza* and for the copies of texts by Amélia Mingas that were unavailable in Portugal; Miguel Oliveira, for keeping me company when I was sick and helping me with the recordings, and Romain Gillain for lending me the recorder and microphone that made these recordings possible; The Fundação para a Ciência e Tecnologia for awarding me the masters grant SFRH/BM/13471/2003 which provided me with the necessary funding to conduct the library research for this thesis and to support me while I wrote it; all the anonymous informants in Angola, for their willingness to cooperate with my research, but above all for showing me that it was relevant to them; and finally, my family, especially Mara and João, for having put up with a workaholic and for helping me cope with the consequences of being one.

# List of Tables


| | | | |
|---|---|---|---|
| *Table* | 1 | Ethnolinguistic groups in Angola (Bantu and Non-Bantu) | *page* 2 |
| | 2 | Demographics in the District of Moçâmedes (1854 and 1864) | 79 |
| | 3 | Population of the colony of Angola (1897-1930) | 87 |
| | 4 | The population in the cities of Angola according to race in 1930 | 88 |
| | 5 | Racial composition of the Angolan population (1940-1970) | 90 |
| | 6 | The population in the cities of Angola (1940) | 91 |
| | 7 | The population in the cities of Angola (1950) | 91 |
| | 8 | The population in the cities of Angola (1955) | 92 |
| | 9 | Level of proficiency in Portuguese (rural areas in Angola) | 93 |
| | 10 | Frequency of use of Portuguese (rural areas in Angola) | 94 |
| | 11 | Nominal thematic classes in European Portuguese | 102 |
| | 12 | Adjectival thematic classes in European Portuguese | 102 |
| | 13 | Person markers in European Portuguese | 106 |
| | 14 | Noun class system in Kikongo | 112 |
| | 15 | Noun class system in Kimbundu | 112 |
| | 16 | Noun class system in Umbundu | 113 |
| | 17 | Noun class system in Cokwe | 113 |
| | 18 | Types of formal number marking in Kikongo | 117 |
| | 19 | Types of formal number marking in Kimbundu | 118 |
| | 20 | Types of formal number marking in Umbundu | 118 |
| | 21 | Types of formal number marking in Cokwe | 119 |
| | 22 | Number marking strategies in Kikongo | 119 |
| | 23 | Number marking strategies in Kimbundu | 120 |
| | 24 | Number marking strategies in Umbundu | 120 |
| | 25 | Number marking strategies in Cokwe | 120 |
| | 26 | Sex-based gender oppositions in Cokwe | 125 |
| | 27 | Free subject markers denoting [+ human] entities in KK, K, U and CC | 126 |
| | 28 | Object person markers denoting [+ human] entities in KK, K, U and CC | 127 |
| | 29 | Reflexive person markers denoting [+ human] entities in KK, K, U and CC | 127 |
| | 30 | Emphatic person markers denoting [+ human] entities in KK, K, U and CC | 128 |
| | 31 | Possessive determiners in KK, K, U and CC | 129 |
| | 32 | Genitive particles in KK, K, U and CC | 130 |

# List of Maps and Figures


| | | | |
|---|---|---|---|
| *Map* 1 | Angola | *page* | 5 |
| 2 | Population density of Angola | | 6 |
| 3 | Bantu ethnolinguistic groups in Angola | | 7 |
| 4 | Central Africa, c. 1500 | | 63 |
| 5 | Portuguese exploration journeys in Central Africa in late 18th century | | 76 |
| *Fig.* 1 | The Animacy Hierarchy | | 116 |

# Abbreviations and Symbols [2]

| | |
|---|---|
| AC | Anobón Creole |
| ADJ | adjective |
| AGR | agreement |
| AP | African Portuguese |
| AUX | auxiliary |
| AVP | Angolan Vernacular Portuguese |
| BP | Brazilian Portuguese |
| BVP | Brazilian Vernacular Portuguese |
| C | noun class |
| CC | Cokwe |
| CLF | classifier |
| CPLP | Comunidade de Países de Língua Portuguesa |
| CVC | Cape Verde Creole |
| DEF | definite |
| EP | European Portuguese |
| F | feminine |
| GBC | Guinea Bissau Creole |
| GEN | genitive |
| HP | Helvécia Portuguese |
| IMP | imperative |
| K | Kimbundu |
| KK | Kikongo |
| L1 | first language |
| L2 | second language |
| LP | Luanda Portuguese |
| M | masculine |
| MP | Mozambican Portuguese |
| NP | noun phrase |

[2] Abbreviations and symbols used in glosses are those of *The Leipzig Glossing Rules*, which have been developed jointly by the Department of Linguistics of the Max Planck Institute for Evolutionary Anthropology (Bernard Comrie, Martin Haspelmath) and by the Department of Linguistics of the University of Leipzig (Balthasar Bickel). A PDF version of these rules is available on the World Wide Web at http://www.eva.mpg.de/lingua/files/morpheme.html

| | |
|---|---|
| PL | plural |
| POSS | possessive |
| PRS | present |
| PRV | partially restructured vernacular |
| PST | past |
| SEP | Standard European Portuguese |
| SG | singular |
| TL | target language |
| U | Umbundu |
| . | neither formally nor semantically segmentable element |
| - | morpheme boundary |
| * | ungrammatical utterance |
| ∅ | zero marker |
| < > | infix |
| 1 | first person |
| 2 | second person |
| 3 | third person |

# Abstract

Building on the theoretical model for varying degrees of restructuring advanced in Holm (2004), this thesis seeks to determine how the morphosyntax of Angolan Vernacular Portuguese (AVP) indicates the degree of restructuring that Portuguese underwent in Angola, i.e. the extent of the retention of its morphosyntax as opposed to the introduction of substrate/interlanguage features. Hence, this thesis surveys the different ways in which Angolan Vernacular Portuguese has been described in the literature (Chapter 2) and provides an in-depth analysis of the relevant sociolinguistic and historical factors shaping its genesis and development (Chapter 3) as well as the first systematic description of the variety's noun phrase morphology (Chapter 4). Hopefully, this will cast much needed light on our understanding of the continuum of contact-induced language changes that has led to the emergence of contemporary AVP (Chapter 5).

# Resumo

Partindo do modelo teórico proposto por Holm (2004) para os diferentes graus de reestruturação linguística, avalia-se em que medida a morfossintaxe do português vernáculo de Angola indica o grau de reestruturação sofrido pelo português nesse país, i.e. o índice de retenção da sua morfossintaxe por oposição à introdução de características de substrato. Assim, esta tese analiza as principais descrições do português vernáculo de Angola disponíveis na literatura (Capítulo 2), descreve os factores sociolinguísticos e históricos relevantes para o desenvolvimento e consolidação da variedade (Capítulo 3) e fornece a primeira descrição sistemática da morfossintaxe do seu sintagma nominal (Capítulo 4). Espera-se, assim, alcançar um necessário entendimento do contínuo das mudanças linguísticas induzidas por contacto que levaram ao surgimento do português vernáculo de Angola contemporâneo (Capítulo 5).

# Chapter 1

## Introduction

### 1.0. Introduction

Angola is located on the west coast of Central Africa on the South Atlantic Ocean, between Namibia and the Democratic Republic of the Congo. The country covers an area of 1,246,700 square kilometers and is divided into 18 provinces: Cabinda, Zaire, Uíge, Bengo, Luanda, Kwanza Norte, Kwanza Sul, Malanje, Lunda Norte, Lunda Sul, Benguela, Huambo, Bié, Moxico, Namibe, Huíla, Cunene, Cuando-Cubango (cf. Map 1, page 5). The population is estimated at roughly 14,000,000 (World Bank 2004), 36% of which is concentrated in the biggest urban centers in the western half of the country (*ibid*), as shown in Map 2 (page 6). Luanda is the capital and the largest city in the country.

While Angola is in size the second largest speaking-Portuguese country, Portuguese is spoken natively by less than 20% of the population, mainly the elite and the young in large urban centers, typically on the coast (Cuesta 1990:15). The remaining 80% of the population, concentrated in the poor neighborhoods of the coastal cities or scattered throughout the rural interior, is monolingual or bilingual in the different African languages spoken in the country. This population belongs to different ethnolinguistic groups, which fall into two groups: the Bantu and the Khoisan (for the geographical distribution of these groups see Map 3, page 7).

As far as the Bantu ethnolinguistic groups are concerned, the names of the people and that of the language derive from a common stem but they differ in the meaning of the prefix that is added to it, as the Bantu languages have specific prefixes to refer to people as opposed to languages (Chatelain 1888-1889:xi-xii). Since the names of the people and of the language are often confused, the two forms are given below in Table 1. Throughout this thesis, the distinction between language and people will be maintained.

**Table 1: Ethnolinguistic groups in Angola (Bantu and Non-Bantu)**
(based on Fernandes & Ntondo:23-56)

| | PEOPLE | LANGUAGE |
|---|---|---|
| BANTU | Tucokwe | Cokwe |
| | Ambundu | Kimbundu |
| | Bakongo | Kikongo |
| | Vangangela | Ngangela |
| | Ovanyaneka-Nkhumbi | Olunyaneka |
| | Ovahelelo | Oshihelelo |
| | Ovambo | Ovakwanyama |
| | Ovandonga | Oshindonga |
| | Ovimbundu | Umbundu |
| NON-BANTU | Khoisan | Khoisan |
| | Vátwa | Vátwa |

The most widely spoken languages are the Bantu languages, six of which were selected in the late 1970s as national languages: Kikongo, Kimbundu, Cokwe, Umbundu, Mbunda-Ngangela, and Ovakwanyama (Instituto Nacional de Línguas 1980). The exact number of speakers of these languages is uncertain, but Bonvini (1994:128) provides the following figures: Umbundu (3,200,000), Kimbundu (2,000,000), Cokwe (1,300,00), Kikongo (1,000,000) and Ovakwanyama (200,000).

The maintenance of the African languages is noticeable not only in the interior, but also in the large urban centers. In fact, according to the census of 1983, 60% of the inhabitants of Luanda were native speakers of these languages (Amaral 1987b:53). However, as a consequence of the fact that Portuguese is both the official language and a vehicular language, the language of the administration, of the media and education, Bantu-speaking Angolans usually have different degrees of proficiency in a variety of Portuguese which has been referred to in the literature as "português angolanizado" (Macedo 2002; Mendes 1985), "português angolano" (Marques 1983) or, more recently, "português vernacularizado" (Fernandes & Ntondo 2002). It is this vernacular, that is both first and second language of a considerable number of Angolans, that the present author refers to as Angolan Vernacular Portuguese. Mingas (1998:115) defines it as follows:

> ... uma nova realidade linguística em Angola, a que chamamos "português de Angola" ou "angolano", à semelhança do que aconteceu ao brasileiro ou ao crioulo. Embora em estado embrionário, o "angolano" apresenta já especificidades próprias [...] Pensamos que, no nosso país, o "português de Angola" sobrepor-se-á ao "português padrão" como língua segunda dos Angolanos.

The first references to this variety in the literature appeared in the late 1980s. Back then, it was not perceived as a distinctively Angolan variety of Portuguese, but rather as a more or less systematic set of deviations from the norm of European Portuguese that could be found in the speech of second-language learners of Portuguese as a consequence of interference from their Bantu mother tongues.

However, as stated by Mingas (*ibid.*), the general perception today among Angolan linguists is that this contact-influenced vernacular is developing into a full-fledged variety of Portuguese, as some of its distinctive features have been incorporated into the speech of most native speakers of Portuguese in Angola. This new approach to AVP is perfectly illustrated by Cabral's assertion below (2005:2):

> A identificação e a descrição de tendências de mudança no Português em Angola devem (pre)ocupar professores e linguistas, podendo estes alertar para a necessidade de uma tomada de posição, por parte do governo, de modo a caminharmos para o estabelecimento da norma culta do Português em Angola, distinguindo aquilo que são construções transitórias, ou seja, próprias de processos de aprendizagem das línguas, daquelas passíveis de serem encontradas na produção de falantes adultos cultos.

The continuous state of warfare into which Angola was plunged from 1975 until recently made it difficult for linguists to collect linguistic data that could be used to develop a comprehensive description of this contact-influenced vernacular. Consequently, the few studies published so far consist mostly of brief conference papers by Angolan linguists, with the exception of Mendes (1985), Mingas (2000) and Cabral (2005). In spite of their scientific quality and valuable contribution to our knowledge of the variety, these studies tend to be short collections of unrelated contact-induced features of AVP. Some of those features include neologisms and lexical borrowings from the Bantu languages, significant changes in the word order of possessive adjectives and person markers, variable number and gender agreement in the noun phrase and between the subject and the verb, significant changes in the lexico-semantic properties of some verbs and in the inventory of EP vowels and consonants.

The data used in descriptions of AVP up to now were exclusively drawn from written texts, rather than from the everyday speech and are often repeated in the works of different linguists, as noted by Gonçalves (2004:229):

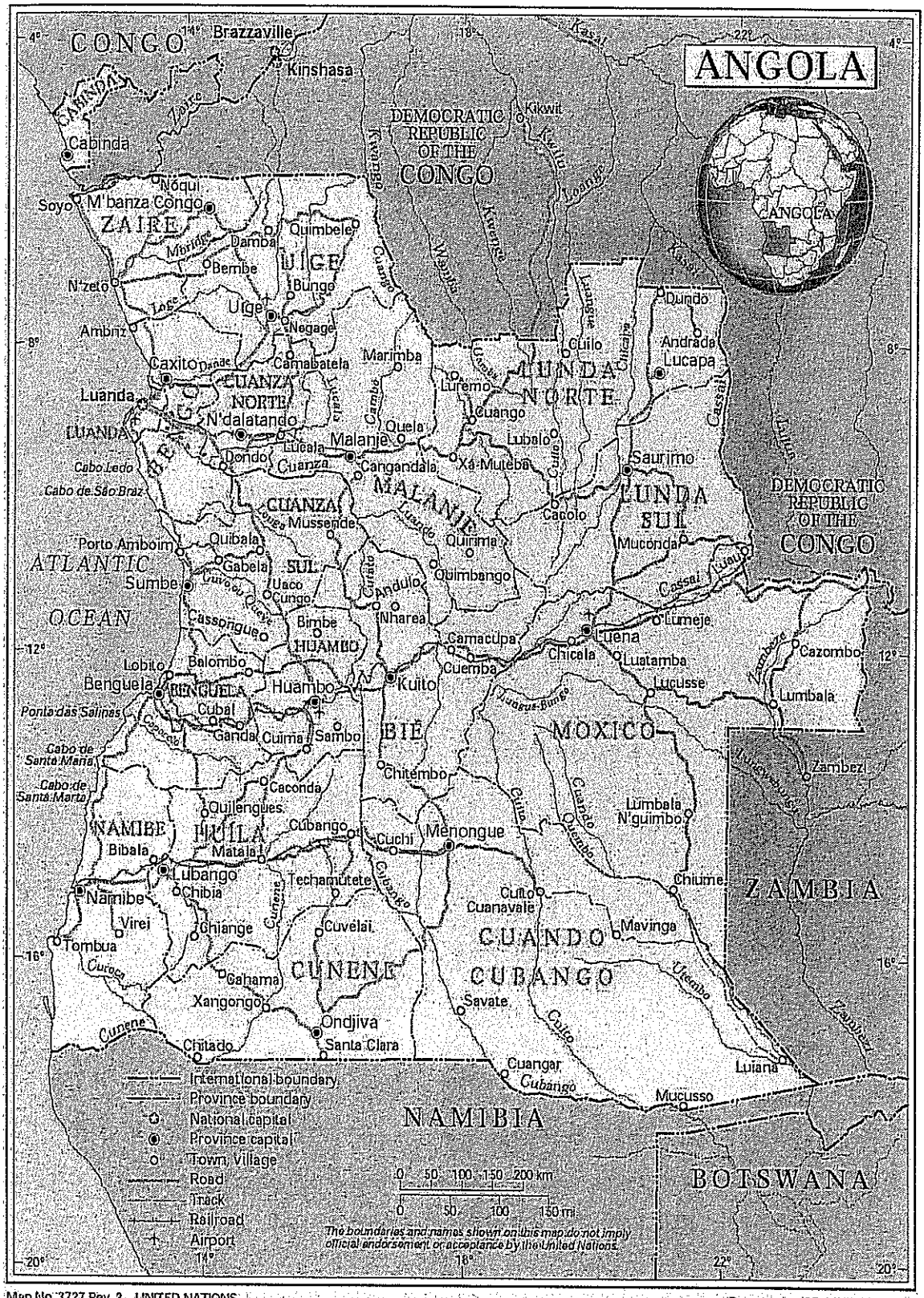
> ... up to now, the studies on this African variety of Portuguese have only been based on texts by Angolan writers, and not on data produced in natural contexts by L2 speakers of this speech community (cf. Gärtner 1996; Marques 1985; Mendes 1985; Mingas 2000). Furthermore, some of these writers (e.g. Luandino Vieira or Manuel Rui Monteiro) are L1 speakers of Portuguese, making it difficult to take their discourse as reliable evidence about the process of L2 acquisition of Portuguese in Angola.

In fact, to the best of the knowledge of the present author, Cabral (2005) and the present thesis are the only sources of oral data produced by actual speakers of the variety.

The most significant feature of the descriptions of AVP published up to now is the fact that they do not analyze it from the perspective of contemporary contact linguistics. Consequently, very little attention is given to the detailed study of the sociolinguistic setting in Angola during the period of contact, despite the fact that this is widely acknowledged as a crucial factor among contact linguists. This has led to several shortcomings. First, no distinction is made between contemporary contact-induced features and those features that developed during earlier stages of contact. Second, it is not possible to determine whether the degree of restructuring of Portuguese in the cities (where its influence is more noticeable) and in the interior (where the majority of the population is still monolingual or bilingual in Bantu languages) is comparable. Third, it is not clear whether other restructured varieties of Portuguese or Bantu languages developed in Angola that may have served as substrate or adstrate languages of AVP. Above all, it is not clear from the existing literature which linguistic and sociolinguistic factors were relevant in the restructuring of Portuguese in Angola or how these combined to produce the specific linguistic features of AVP.

Nonetheless, in spite of the these shortcomings, the existing literature on AVP makes it clear that the vernacular that is spreading throughout Angola as either the first or second language of the general population owes its structure to the contact-induced restructuring of Portuguese in the country. Furthermore, it shows that the features that distinguish this vernacular from European Portuguese bring it closer to other partially restructured varieties such as Brazilian Vernacular Portuguese, in that it displays considerable influence from substrate languages, but it also retains enough similarity to the superstrate to indicate that it has never been fully creolized.

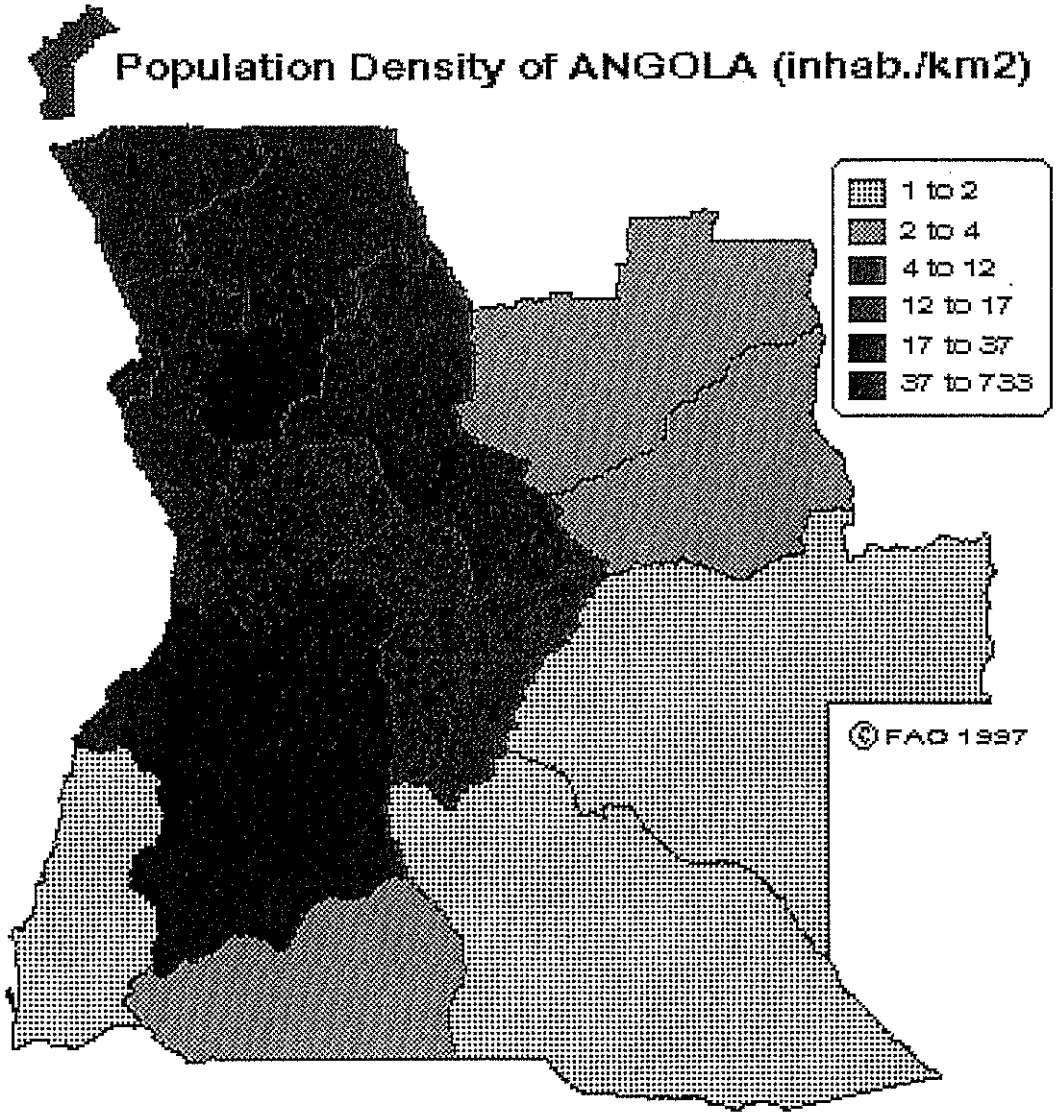
**Map 1: Angola**
(United Nations 1997)



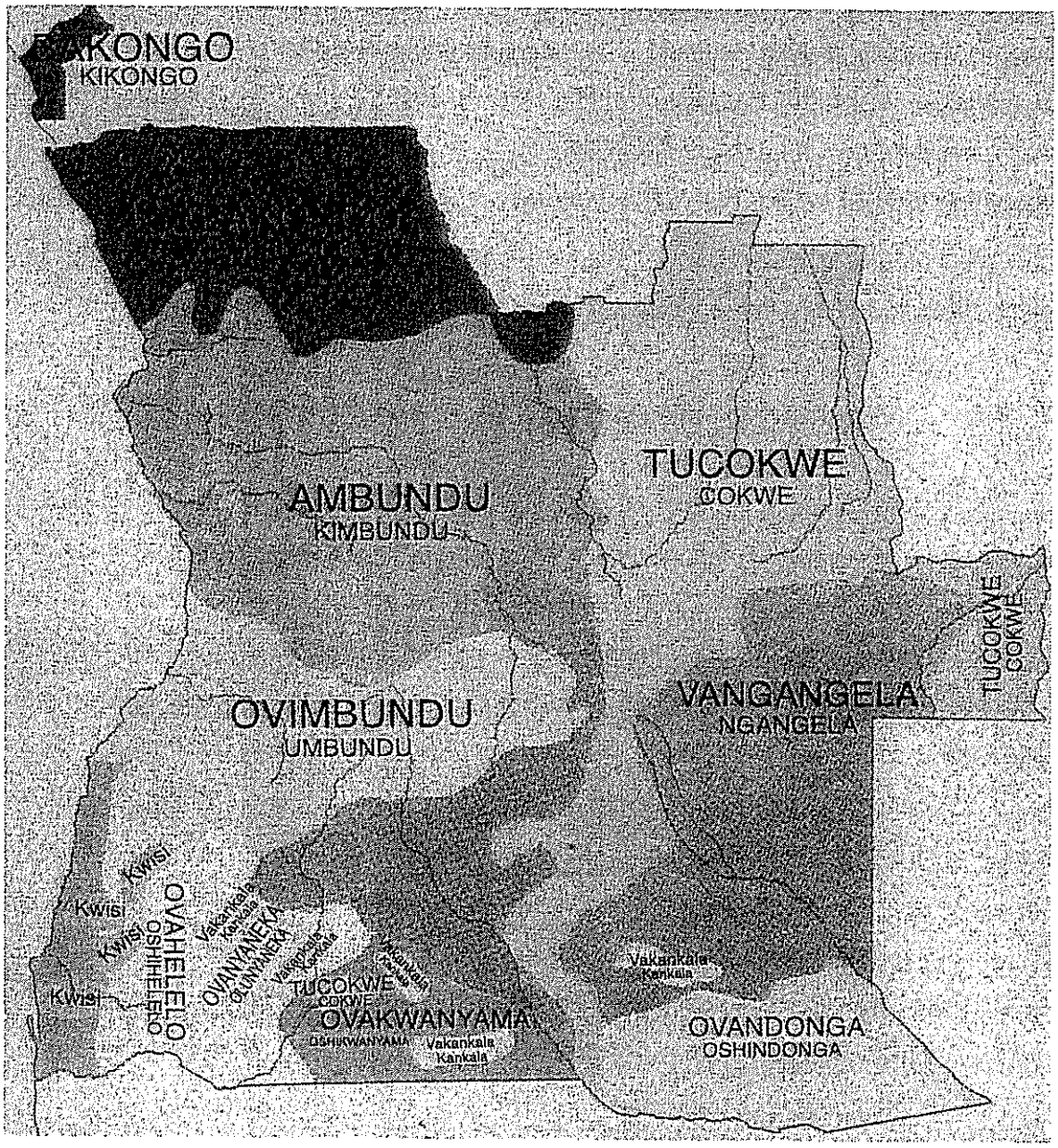

Map No. 3727 Rev. 2 UNITED NATIONS
October 1997 (Colour)

Department of Public Information
Cartographic Section

**Map 2: Population density of Angola**
(FAO 1997)



**Map 3: Bantu Ethnolinguistic groups in Angola**
(Fernandes & Ntondo 2002:57)




## 1.1. Object of study and aims of the thesis

This thesis is intended to be a preliminary contribution to a more comprehensive understanding of the historical development and synchronic structure of the vernacular Portuguese that has arisen in Angola, rather than a quest for a single theoretical model to account for it. Hence, building on the recent contributions from contact linguistics, especially Holm's (2004) model for partial restructuring, the present thesis focuses on a description of the morphosyntax of the noun

phrase in AVP with the purpose of determining how this might indicate the degree of restructuring that Portuguese underwent in Angola. Four morphosyntactic phenomena are analyzed: number and gender marking as well as agreement, possession, word order and case marking of object person markers. The aim is to understand these findings within the sociolinguistic context of language contact in Angola from the early days of settlement to its independence from Portugal in 1975 to determine how Portuguese came to be restructured and developed into a vernacular language in Angola.

Thus the main goal of the present thesis is to identify the social and linguistic factors that have shaped the historical development and synchronic structure of AVP.

## 1.2. Structure of the thesis

In order to achieve this goal, this thesis includes four further chapters. Chapter 2 seeks to provide an overview and critical evaluation of the most relevant literature on both AVP (2.1) and contact linguistics (2.2). While it does not provide an exhaustive review of all the existing literature on these topics, it tries to provide an insightful outline of the different ways in which AVP has been studied and described, and to focus on the current debates within contact linguistics.

Chapter 3 analyzes the sociolinguistic setting in Angola from the arrival of the Portuguese to the present. It identifies six main periods in the history of the Portuguese presence in the country and it evaluates how social factors such as the demographics of the populations in contact, the level of bilingualism, the displacement of the substrate speakers, the degree of learners' access to the target language, as well as the social patterns of colonization affected the nature and extent of language contact in each of those periods.

Chapter 4 evaluates the influence of the structure of the relevant languages in Angola in the restructuring of AVP's noun phrase concerning the morphosyntactic phenomena discussed above. Section 4.2 focuses not only on the description of the noun phrase in KiKongo, Kimbundu, Umbundu and Cokwe, but also on the description of noun phrase morphology of the Bantu languages in general. This is due to the importance of clearly understanding the complex system of nominal classification and agreement displayed by all Bantu languages. In fact, while this system displays language-specific features, these cannot be understood without previous knowledge of the general features that define it. Hence, prior to the analysis of the four morphosyntactic phenomena in Kikongo, Kimbundu, Umbundu and Cokwe, a general description of the Bantu nominal classification system is provided. Moreover, this general

description of the NP morphology in Bantu may indicate grammatical features that are not displayed by the Bantu languages surveyed but which may have played a role in the restructuring of AVP, for Kikongo, Kimbundu, Umbundu and Cokwe were not the only Bantu languages with which EP was in contact in Angola. Finally, on purely methodological grounds, it would be impossible to analyze number and gender marking, possession and word order and case of personal pronouns outside the context of this system of nominal classification.

Finally, chapter 5 relates the findings in the three preceding chapters and attempts to provide a tentative hypothesis concerning the social and linguistic factors that were relevant in the historical development that led to the synchronic structure of AVP.

Examples, tables, maps and figures are numbered consecutively throughout the chapters, but footnotes are numbered individually for each chapter. Any symbols or abbreviations in this thesis, whose consecutive use is avoided as much as possible, are given in full when they first occur. Nonetheless, they are compiled on page v. The original language of quotations was maintained, not only because this best depicts the ideas of their authors, but also because this thesis was written for a readership likely to be familiar with both Portuguese and English, as well as French.

## 1.3. The data

The linguistic data used in this thesis refers to Angolan Vernacular Portuguese, European Portuguese and the Bantu languages with which AVP was in closest contact in Angola (i.e. Kikongo, Kimbundu, Umbundu and Cokwe). The data on European Portuguese is mostly drawn from reference grammars of the language, as well as the author's native competence.

The data on AVP was initially intended to be drawn from the existing literature; however, the author was able to spend two months doing fieldwork in Angola, making possible the collection of a considerable corpus of data on the variety of the vernacular spoken in Dundo (Province of Lunda Norte) which was included in the thesis as a means of supplementing the exclusively written data in the literature with everyday productions of AVP speakers (all data not otherwise identified in this thesis). The data were collected by means of interviews with people of different ages, sexes and levels of instruction, following the Ethical Guidelines of the Department of Linguistics at the Max Planck Institute for Evolutionary Anthropology. Whenever this information is thought to be relevant, it is indicated in the text. The transcription of the data followed two simple guidelines: (i) don't type anything that isn't there, and (ii) do include everything that is there.

Concerning the data on Kikongo, Kimbundu, Umbundu and Cokwe, three preliminary remarks are in order. The first concerns the designations used here to refer to these languages, as a common problem in the existing literature on African languages is the use of different names to refer to a single language, which makes it difficult for the non-native linguist to decide whether a particular work refers to the language one is analyzing or to a completely different language. For example, *Kikongo* is often referred to as Congo, Kisikongo, Kikoongo, Kongo San Salvador. Alternate names for *Kimbundu* include Loanda Mbundu, Kimbundo, North Mbundu, Nbundu, N'bundo, Dongo, Ndongo, Kindongo. Those for *Umbundu* include Umbundo, M'bundo, Ovimbundu, South Mbundu, Mbundu Benguella. *Cokwe* is some times referred to as Chokwe, Ciokwe, Shioko, Kioko, Quioco, Djok, Tshokwe, Tschiokloe (Gordon 2005).

There are several reasons for this: (a) the fact that different researchers use different spelling conventions or hear different things depending on their own language backgrounds; (b) the fact that different forms of the same name sometimes derive from dialectal differences within the language, or between the language and closely related languages; (c) or the fact that the names of dialects of language have been extended to the language as a whole (cf. Wald 1994:289-90). In the Angolan case, it must also be taken into account that the name of the people has sometimes been extended to the language (e.g. *Ovimbundu* is often used to refer to *Umbundu*). Hence, the designations used in this thesis are those suggested by the Angolan National Commission of the International Institute of the Portuguese Language (cf. Fernandes & Ntondo 2002), i.e. Kikongo (KK), Kimbundu (K), Umbundu (U) and Cokwe (CC).

The second preliminary remark concerns the nature of the data drawn from the existing literature, which includes 19th century grammars written by missionaries with little or no specific linguistic training. Of course, this can present several problems. First, the lack of specific linguistic training of authors requires a cautious use of their analyses. Second, assuming that these are accurate, it is possible that at least part of the data they present no longer conforms to the contemporary use of the languages. Hence, the analyses and data they provided in these earlier grammars are presented along with recent descriptions of the relevant Bantu languages.

The orthography used in the examples on KK, K, U and CC follows that of the sources, but morpheme boundaries have been added in order to facilitate interpretation of the data. The examples are glossed according to the Leipzig Glossing Rules.

# Chapter 2

## Overview of the literature on AVP and Contact Linguistics

### 2.0. Introduction

The goal of this chapter is threefold. First, section 2.1 seeks to summarize and critically evaluate the most relevant information concerning the use of Portuguese in Angola in order to provide an outline of the different ways in which it has been studied and described, which will hopefully cast some light on the ways Portuguese has developed there from the earliest contacts to the present. Second, section 2.2 aims at presenting the theoretical foundations of this thesis by providing a brief overview of the most recent literature on contact linguistics and a critical analysis of the current debates within the field. Finally, section 2.3 builds on the information provided in sections 2.1 and 2.2 to formulate a working hypothesis to account for the origin of the synchronic structure of AVP.

### 2.1. The study of Angolan Vernacular Portuguese

While the description and classification of the Portuguese-based creoles in Africa goes back to the nineteenth century and a growing body of work is being developed on the restructured variety of Mozambican Portuguese, i.e. MP (Gonçalves 1985, 1990, 1998, 2004; Stroud & Gonçalves 1997), little has been done regarding the description of the development and structure of AVP.

The civil war that devastated the country since its independence from Portugal in 1975 until recently is responsible for this, in the sense that it made fieldwork in Angola difficult and dangerous, but the Portuguese neo-colonial myth of *Lusofonia*[1] is no less to blame, fostering a

[1] The word *Lusofonia* refers to the community of countries where Portuguese is used either as a first or a second language, i.e. Portugal, Brazil, Cape Verde, Guinea Bissau, São Tomé e Príncipe, Angola, Mozambique and East Timor. According to Pérez (2000:1), *Lusofonia* is simultaneously a common linguistic community and economic area, a political organization embodied by the CPLP (i.e. Comunidade de Países de Língua Portuguesa ); and a cultural organization achieved by a series of cultural agreements.

sentimental and false view of the country as exclusively Portuguese speaking (Lopes 2003), even if some differences between AVP and European Portuguese (EP) are recognized. This view has not led to investing in the study of actual language use in Angola (Léonard 1995; Pérez 2000) and there are very few detailed and systematic linguistic studies of AVP. The following sections provide a brief overview of the most relevant of these studies, although some of them contain no more than scattered references to Portuguese language use in Angola.

### 2.1.1. Reference works in linguistics

If one looks under the entries *Portugal* or *Portuguese* in most reference works in linguistics, one is left with the impression that there is no such thing as an Angolan variety of Portuguese. On the one hand, this is because Portuguese is described as an extremely homogeneous language and, on the other hand, because only two varieties of Portuguese are acknowledged, i.e. EP and Brazilian Portuguese (BP). Castilho (1994:3234) is a case in point. Apparently, he considers that the only difference between the Portuguese spoken in Portugal and Brazil and the Portuguese spoken in Africa is its socio-political status, i.e. native language as opposed to official state language, respectively. No African variety of Portuguese is referred to under the sections on the characteristics and varieties of Portuguese. This indicates that he believes that they are in no way different from EP, which he considers to present "a notable lack of differentiation, with the variety of Lisbon providing the standard" (*ibid.* 3235).

Dictionary and encyclopedia entries on *Angola* are usually more enlightening, despite the fact that, for example, the estimated numbers of Portuguese speakers in the country are usually too low and imprecise. For example, Grimes (1996) estimates there are about 57 600 Portuguese speakers in Angola. However, not only does she not distinguish between native and non-native speakers but she also clearly overlooks the growing number of Angolans, especially those born in the cities to Bantu-speaking refugee families, who, due to the pressure of schooling and the media, now primarily speak Portuguese (both first and second language).

---

The truth though is that after the independence of Portugal's ex-colonies, the concept of *Lusofonia* came to replace the colonial myth of taking these countries as solely Portuguese speaking, thus obscuring the fact that, apart from Portugal and Brazil, the majority of the population in the "países lusófonos" speaks Portuguese as second or third language rather than a first language. The general low academic success of students from the PALOPs (Países Africanos de Língua Oficial Portuguesa) in Portugal clearly illustrates the dramatic consequences of this myth in people's lives, i.e. they are assumed to be native speakers of Portuguese and are evaluated as such. As Léonard (1995:16) puts it "Le mythe est en marche". For a short but incisive and recent description of people's attitudes towards the minority languages in Portugal see Pereira (2002).

Still, the description of language use in Angola provided in these entries is often more realistic than that of Portuguese provided by Castilho (1994). Bonvini (1994:127-28) is a case in point:

> Portuguese was chosen as and confirmed the official language when Angola became independent in 1975, after having been introduced in the country in the fifteenth century by the first explorers, used as lingua franca for the slave trade and commerce, imposed as civilization language by the colonizers, and adopted as a prestige language by the 'civilized' (assimilados). In the last decade of the twentieth century it is used in all areas of public life (administration, mass media, education, and international relations) and spoken in particular by younger people and the elite especially in the bigger towns of the coastal area. However, it must be pointed out that through contact with the local African languages, especially in the rural areas, it has undergone significant changes both lexically and syntactically. As a result it gave rise to social variants, the most typical of which being the *linguagem dos musseques* in Luanda, which however did not reach the status of regional language specific to Angola.

Bonvini is not so much concerned with the socio-political status of Portuguese in Angola, which has continuously changed over the years, i.e. lingua franca, civilization language, prestige language and official language, but with the past and ongoing consequences of those different statuses, i.e. an extremely complex sociolinguistic situation in which a distinction must be made between L1 Portuguese, spoken by the young and the elite, mainly in the coastal cities, and L2 Portuguese, which Bonvini refers to as "contact Portuguese" or "social variants" of Portuguese, spoken especially in the interior and in the poorest urban neighborhoods, e.g. *musseques.*

While Bonvini (1994) is careful about acknowledging the existence of "a regional language specific to Angola", his description of language contact in the country indicates that in recent years the sociolinguistic conditions for such a variety to emerge have been met:

> Various factors such as slavery, commerce, colonial wars and wars of independence have caused a permanent linguistic mix between Portuguese and the main African languages. Therefore the number of two-way borrowings is increasing, and there is a clearly diglossic situation[2] in coastal areas and towns. In the last decade of the twentieth century this situation has been reinforced through unprecedented massive migrations and urbanization owing to pre- and post-colonial wars. (*ibid.* 128)

It is the description of AVP as a language variety in the making that is remarkable in dictionary and encyclopedic entries on Angola, especially when compared with the descriptions made by most Portuguese linguists writing before the mid-20th century, for whom AVP simply does not exist.

---

[2] My fieldwork in Angola showed that there is some degree of diglossia in the urban centers, but it seems to be more characteristic of the rural areas, where most children and elders speak one or more indigenous languages as their first languages and use Portuguese only in formal contexts, e.g. school, or in contexts where this language is perceived as the norm, e.g. the center of the urbanized areas.

### 2.1.2. General works about varieties of Portuguese around the world

While the accounts of early travelers provide occasional anecdotal references to the use of Portuguese outside of Portugal, the first general descriptive Portuguese studies of the language as spoken in areas other than Europe go back to the end of the nineteenth century only.

The main emphasis of these early works was on the description of Portuguese-based pidgins and creoles. The works of Adolfo Coelho (1880-86) are paradigmatic in this regard. All three papers are devoted to the description of the creoles based on Portuguese and on other European languages in Africa, Asia and the Americas. Of the non-creole varieties of Portuguese, only BP is analysed, classified as "um português diferenciado [...] onde se manifesta uma tendência crioulizante" (Coelho 1886:160,170). Coelho may have been "the first to articulate a theoretical position on the origin of creoles which came to be called the universalist theory" (Holm 2000:27), but as far as AVP is concerned, he provided no information whatsoever.

The first description of AVP came from abroad, in the work of Hugo Schuchardt, who in 1888 published "On Creole Portuguese", a paper mainly devoted to the description of the "mainland creole" (Schuchardt 1888:60), spoken on the Gold and Slave Coasts, Angola and Mozambique. Schuchardt based his work on data and remarks collected from the reports of people living in those areas, most of whom described the Portuguese spoken in Angola as "quite corrupted" (*ibid.* 67). This has been interpreted by some as evidence for the existence of a possible Portuguese pidgin or creole in Angola from which the present vernacular may have sprung. The most quoted of these remarks dates from 1886 and is that of Macedo Soares:

> Dá-se mesmo um facto curioso em Angola e mais possessões portuguezas d'Africa Austral: a coexistência de três vocabulários: o *portuguez* fallado pelos portuguezes entre si; o *bundo*, pelos negros entre si; e um intermédio, a que chamaremos *mestiço* ou *crioulo*, usado nas relações dos negros com os brancos, e também pelos estrangeiros quando se querem entender com portuguezes ou com negros. O mestiço se compõe de palavras portuguezas accomodadas ao génio do bundo e tende a se generalizar e firmar, por isso mesmo que é percebido e fallado pelos três grupos da população: negros, portuguezes e estrangeiros. (Schuchardt 1888:67 quoting Soares 1886:14)

This remark tells us that the Portuguese and blacks in Angola used at least three different vocabularies, but not necessarily three completely distinct languages, i.e. Portuguese, Kimbundu and an intermediate language. Interpreting this as evidence for the existence of a Portuguese-based creole in Angola is probably a mistake. Three facts seem to support this hypothesis. First, Soares's description testifies to the existence of widespread bilingualism in Angola by the end of the 19th century, which generally inhibits the emergence of creoles. Second, the use of

Portuguese words accommodated to the morphosyntax, i.e. "génio", of Kimbundu does not necessarily imply the forging of a new language, as I believe is shown by the presence of both lexical and grammatical Portuguese words in the following examples in Kimbundu[3]:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (1) | O | mama | inga | utuma | kuxitisa | o | mabitu | ni | jinjanena |
| | The | mother | then | orders | to block | the | doors | and | windows |

(Chatelain 1894:30)

| | | | | |
|---|---|---|---|---|
| (2) | Fenda Madía | uiu | bu | _njanena. |
| | Fenda Maria | goes | to | the window (*ibid.* 35) |

Finally, it is Schuchardt (1888:64) himself who states that:

> ... as far as the quality of the Portuguese ... is concerned, to be sure, there is often mention of one aspect of it in travelogues that raises the hope of fruitful insight. For example, Oberländer (1874:221) says of Weida[4] that: "the numerous mulattos occupy a specific quarter of the city and speak a sort of Portuguese jargon." However, in my personal research I have not been able to uncover the slightest example of such a jargon. An Englishman transcribed an anecdote for me in the Portuguese that is spoken there, but it is Standard Portuguese.

If to this we add that "Portuguese gained a firmer hold in Lower Guinea than in Upper Guinea" (*ibid.* 65) it is even more difficult to accept the existence of a fully developed pidgin or creole in Angola as late as the 19th century, even if there were varying degrees of interference between Portuguese and the Bantu languages, depending on the geographic location of the speakers:

> If the Blacks speak Portuguese most correctly in the main cities of the province where schooling is available to them, then they speak it less perfectly in those areas where it serves them in trade with other Europeans. This is primarily the case in the area north of Ambriz. (Schuchardt 1888:68)

Schuchardt (1888:68-71) later on elaborates on what he means by Bantu speakers using Portuguese "less perfectly", but never does he conclude absolute regularity:

> Of the phonetic features, the almost regular insertion of a vowel in words terminating in *l* or *r*" is to be underlined, e.g. *mulhera, favoro, liquoro, papelo, solo* [...] It is with a certain comic regularity that *r* and *l* are interchanged for one another in Portuguese words, e.g: *rolla* 'dove' becomes *lora*". (*ibid.* 68)

As far as morphosyntax is concerned, the situation was also not completely stable yet, for "while inflection has become extremely restricted. To be sure, the sigmatic plural[5] is in use,

---

[3] Examples of both lexical and grammatical Portuguese words incorporated into the grammar of Olunyaneka, a Bantu language spoken in the province of Wíla (Fernandes & Ntondo 2002:49), can be found in Estermann (1971).

[4] Ajudá (in Benin).

[5] i.e. plural expressed by an inflection.

though not after numerals: *tres boi, dez casa*" (Schuchardt 1888:69). In a nutshell, examples of zero plural markers coexisted side by side with full inflectional plural marking.

The verb morphology of AVP had also undergone restructuring. Schuchardt (1888) notes that "In the personal endings of the verbs, the third person ending stands for the other desinences: *eu tem, eu está, eu vae*" (*ibid.* 69) and that "the perfective (completely and definitely terminated time past) is almost never used, only the imperfective" (*ibid.* 70). However, while he considers these features typical of a "germinating creole" (*ibid.* 69), he notes that "inflected tenses have not died out" in AVP and therefore the occurrence of more creole-like features, i.e. the use of infinitives instead of finite verbs without preceding temporal markers, remains to be attested.

In short, even if we bear in mind that Schuchardt was working on incomplete and indirect data, the overall description of AVP provided by him appears to be one of a language variety in formation, not one of a fully developed pidgin or creole out of which AVP could have sprung. This hypothesis may be further supported if one takes into account that some of the features of AVP reported by Schuchardt may have developed elsewhere, on the other side of the Atlantic, in Brazil. Schuchardt's following remark about the survival of Portuguese on the coastline is elucidating:

> ... there was also Portuguese occupation. But the Portuguese were too poorly protected by the dilapidated fort at Ajudá, or, more to the point, by its weak garrison which consisted mainly of deportees and Luanda Creoles. The fact that the slave trade was very lively and conducted most recently here is more significant. For this reason, as well as the peaceful conduct of the copra trade, the petty kings, their nobility, and their officials cultivated Portuguese. The one essential fact is the return of the vast number of emancipated slaves from Brazil to these regions, to Porto Seguro, both Popos, Agué, Weida, Godomé, Cotonou, Porto-Novo, Badagry, Lagos, etc. (Schuchardt 1888:62-3)

The early 20th century would see the birth of Portuguese dialectology. However, this did not lead to any studies about the use of Portuguese in Angola. For example, while José Leite de Vasconcellos, who is unanimously recognized as the father of Portuguese scientific dialectology, devoted his 1901 PhD dissertation at the University of Paris to the description of all Portuguese varieties, both creole and non-creole, he makes no more than a few brief references to the description of AVP.

According to Vasconcellos (1901:27-9), Portuguese was composed of four different groups of "dialects": "dialectes continentaux, dialectes insulaires (i.e. Azores and Madeira), dialectes d'outremer" and "Portugais des Juifs". All restructured varieties of Portuguese, i.e. creole and non-creole, were included in the group of the "dialects d'outremer". It is interesting to note that

both creoles and BP are referred to as "dialects", i.e. "dialectes creoles" and "dialecte brésilien" respectively, but AVP and Mozambican Vernacular Portuguese are merged into a generic grouping, i.e. "Portugais des côtes d'Afrique". It is clear from Vasconcellos' following statement that this was because he did not recognize much relevance to the particularities of those linguistic varieties:

> Les conditions dans lesquelles le portugais s'est implanté dans de si diverses régions d'outremer ont donné lieu tantôt à dês parlers spéciaux, comme les dialectes créoles, tantôt à de simples particularités dialectologiques de moindre importance. (*ibid.* 26)

According to Vasconcellos (1901:26) those "particularités dialectologiques" were basically a matter of lexical borrowing of Bantu into Portuguese and vice-versa, and of slight phonetic changes, e.g. "*ô* pour *ou*, *ê* pour *ei*, devant une consonne et à la fin de la phrase, par ex. *pôc, andô, Janêro, andê*...", which he attributes to the influence of dialects of Southern Portugal (*ibid.* 157).

Based on his contact with an individual from São-Paulo-de-Luanda who had lived for a long time in Moçâmedes, Vasconcellos identifies other "particularities" of AVP, i.e. variation in the word order of pronouns, even in the local newspapers; errors in the use of the Portuguese preposition *em,* and difficulties in distinguishing the Portuguese *r* and *l* (*ibid.* 158). While Vasconcellos recognises that these are features shared with the creoles, he minimizes their importance, even when there are morphosyntactic consequences:

> Parfois, des phénomènes très curieux ont lieu. Comme dans les idiomes bantous, le singulier se distingue du pluriel au moyen de préfixes, et en quimbundo au préfixe *ri-* au singulier correspond *ma-* au pluriel: il arrive dans le mot *machado* «hace», qui, selon les lois phonétiques, est devenu *maxâlu*, les Nègres voient un pluriel formé à l'aide du préfixe *ma-*, et ils lui donnent un singulier *ri-xâlu*. Des faits de cette catégorie, quoiqu'ils soient curieux, n'ont cependant rien d'extraordinaire. (*ibid.* 159)

In fact, as far as AVP is concerned, Vasconcellos' work is relevant not for its detailed description of the vernacular, but for his important insights into the sociolinguistic conditions that could account for its development, and his comparison of it with those varieties from places where creoles and BP developed. In the following statement, for example, he not only points out adult second language acquisition as a determining factor for the restructuring of Portuguese, but also refers to the existence of different degrees of restructuring:

> ... les Portugais ont été obliges d'apprendre quelquefois les langues indigènes, et les indigènes d'apprendre la langue du Portugal. Le second fait est le seul qui m'intéresse pour le moment, parce qu'il en est résulté la formation des dialectes créoles, et d'autres variétés du portugais; entre les uns et les autres, on peut admettre des degrés. (*ibid.* 131-2)

In stating his agenda for Portuguese dialectology, he makes an apology for the use of a descriptive and sociolinguistic approach:

> Donc la DIALECTOLIGIE PORTUGAISE doit être, d'après ce que l'on vient de dire, l'étude non seulement des parlers populaires du Portugal, du galicien et des idiomes de la frontière hispano-portugaise, mais aussi des modifications qu'a éprouvées notre langue dans les régions lointaines où elle a été portée par des conquérants, des colons, ou de simples groupes d'émigrants, et en même temps l'examen des conditions dans lesquelles se sont opérées ces évolutions... (Vasconcellos 1901:32)

Vasconcellos (1901) does not provide a definite classification for AVP, but he does state that while Angola and Mozambique were the two main Portuguese colonies, "aucun dialecte creole ne s'est développé dans ces provinces" (*ibid.* 157). Insufficient as it may seem, his work remains to this day the best overall contribution of Portuguese dialectology to the study of contact varieties of Portuguese; the dialectologists who followed, namely Boléo & Silva (1962) and Cintra (1971), would focus mainly on the description of continental Portuguese dialects. It is true that there was also a significant body of work published on varieties of Portuguese other than EP. However, its purpose was not the description of those varieties but rather the statement of the absolute homogeneity of Portuguese[6]:

> Se pusermos de lado os crioulos pròpriamente ditos ... e que têm características próprias, o que surpreende e causa admiração nesta área tão extensa do português é a sua grande unidade. (Boléo 1955:8-9)

While creoles were recognized as languages on their own right, the whole of Portuguese linguistic diversity was restricted to the only two standards acknowledged, i.e. the one of Portugal and that of Brazil. All other varieties of Portuguese represented no standards of their own. Moreover, while the differences between EP and BP were recognized, they were made synonymous with the lexicon and style only:

> Essa unidade do português torna-se ainda mais flagrante se compararmos as diferenças relativamente pequenas, que existem entre o português europeu e o português do Brasil. [...] a língua portuguesa apresenta uma riqueza e variedade surpreendentes, tanto no que toca ao léxico como no que respeita às linguagens que se podem notar dentro da língua comum... (*ibid.* 9-10)

Interference resulting from contact between Portuguese and other languages was recognized and seen as a source of enrichment for both, but again only lexical borrowing and transfer were perceived as relevant. Therefore, most studies consisted almost entirely of extensive lists of words borrowed from Portuguese into African languages or vice versa (e.g. Martins 1958).

---

[6] This can be inferred from some of the titles of these works, i.e. *Unidade e diversidade da língua portuguesa* (Boléo 1955); *A difusão da língua portuguesa em África* (Carvalho 1971) and *O português entre as línguas do mundo* (Fonseca 1985).

Different outcomes of language contact were acknowledged by some authors (Morais Barbosa 1969) but overall the emphasis was on the description of Portuguese-based creoles, on the systemic unity of Portuguese as opposed to its lexical diversity and on the enforcement of language policies that might achieve the replacement of all indigenous languages by Portuguese. In general, Portuguese linguists gave little thought to the sociolinguistic history of the non-creole varieties of Portuguese in Africa.

The turning point in the study of language contact in Africa in general and of non-creole varieties of Portuguese in particular arrived in the mid 1970s, following the independence of most Portuguese ex-colonies. The emphasis of the works published during this period was still on the description of lexical borrowing and transfer, but unlike most Portuguese dialectologists, foreign linguists felt that a solid description of language contact in Africa implied an in-depth analysis of the sociolinguistic background against which the lexical borrowing and transfer had taken place. It is Bal (1979:13) who sets the agenda for the study of language contact in Africa:

> On devrait ensuite étudier ces phénomènes d'emprunt sous divers aspects, notamment sous un aspect interne: l'adaptation phonétique, morphologique, syntaxique et sémantique des emprunts au système des idiomes emprunteurs, et sous un aspect externe, socioculturel, dans la dimension historique et dans la dimension géographique : la motivation des emprunts, en relation avec des faits historiques, des contacts de cultures ; la délimitation des aires de diffusion des emprunts et la mise en rapport de celles-ci avec l'histoire du commerce et des communications (ports, factoreries, marchés indigènes, pistes de caravanes) ainsi que de la pénétration européenne en Afrique.

Valkhoff (1975) is a clear example of this sociolinguistic approach to the study of Portuguese and African language contact in Africa. However, his greatest contribution to a new approach to the study of non-creole contact varieties of Portuguese was the emphasis on the role played by the Portuguese *lingua franca* spoken along the coasts of Africa from the second half of the 15th century to the end of the 18th century:

> ... no ano de 1974 a língua portuguesa surge como uma *lingua franca*, graças à qual não só os Africanos e os Portugueses, mas também pessoas que falam línguas africanas diferentes, se podem entender. Esta situação já é muito antiga. (...) Esta velha *lingua franca* portuguesa assumiu várias formas mais ou menos crioulizadas segundo os lugares e as circunstâncias. Além deste português crioulo ou *baixo português*, como se dizia em Ceilão, os oficiais holandeses e portugueses, por exemplo, falavam o *português alto* ou literário... (Valkhoff 1975:9-10)

In fact, this idea had already been suggested in Valkhoff (1966), which is actually the very first attempt to account for the origin and synchronic structure of AVP. According to Valkhoff (1966), due to growing contact with standard European Portuguese, in Angola this lingua franca had developed into a variety closer to the standard but with sporadic deviations from it, i.e. a

"secondary creole" (*ibid.* 34). Hence, according to this line of thought, synchronic features of AVP are to be taken as the remnants of this old Portuguese-based pidgin.

Valkhoff (1966) gives almost no linguistic examples other than the accounts of early travelers, most of whom could not speak or understand Portuguese. Still, Valkhoff (1966) was influential in at least three ways. First, he allowed for different degrees of restructuring of the *lingua franca* (depending on the geographical location and historical context) and for its coexistence with standard or literary forms of Portuguese, which opened the door for research into the sociolinguistic backgrounds that could account for these differences. Second, he laid the foundations for a monogenetic theory to account for the origin and development of AVP, i.e. the linguistic traits that distinguish it from EP, like those of most creoles in the world, can be traced to a single source: the Portuguese-based *lingua franca.* The most recent formulations of this hypothesis are Perl (1989) and Lipski (1995)[7]. Finally, he reinforced the tendency to focus on the description of lexical borrowing and transfer, since, to some extent the goal of the works on AVP following Valkhoff (1966) was to uncover the structure of the *lingua franca*:

> Ce qui ne fait pas doute, c'est l'existence et la large diffusion d'un pidgin portugais sur les côtes d'Afrique, à partir du XVI$^{e}$ siècle. Ce pidgin, connu sous le nom de «langue franque» et dit parfois *porto*, est bien attesté. [...] Il nous semble que les si nombreux emprunts d'origine portugaise que se rencontrent ... dans les idiomes les plus divers d'Afrique noire, ne remontent pas au portugais cultivé mais à cette «langue franque». La récolte systématique de ces emprunts, la comparaison avec les parlers créoles portugais d'aujourd'hui et avec la *língua de preto* notée aux XVIe et XVIIe siècles devrait permettre de reconstruire, dans ses grandes lignes, l'ancien pidgin portugais d'Afrique... (Bal 1979:34)

Regardless of the reservations one may have about the role played by this *lingua franca* in the development of AVP[8], one cannot dismiss the relevance of analyzing the phonetic, phonological, and morphosyntactic changes undergone by the extensive corpora of Portuguese loanwords provided by the works published during the 1970s. Concerning number and gender agreement in the noun phrase in AVP, an analysis of this sort could be of great help to our understanding of whether it is the random result of imperfect acquisition of Portuguese, which is the generally accepted theory, or whether it is the continuation of phonological and morphosyntactic mechanisms previously used in a language contact situation.

Oliveira (1990:69-89) illustrates the interest of this type of analysis when referring to the influence of Portuguese on Kimbundu, which was widespread and affected not only the lexicon

[7] The contribution of these authors to our understanding of the development and synchronic structure of AVP will be analyzed with greater detail later on.

[8] A more detailed appraisal of this issue will be given in due place.

(by means of loanwords) and phonology (by means of the adaptation of Portuguese loans to the phonology of Kimbundu, e.g. creating a syllabic pattern of CVCV by adding extra syllables to most words), but also affected the morphology of Portuguese loans to Kimbundu[9]:

> Quanto à morfologia, as perdas das desinências de género e número são consequência da integração dos nomes [i.e. *from Portuguese*] nas classes nominais, inexistente a primeira distinção no Quimbundo, a segunda expressa pela variação dos prefixos de cada classe. Verifica-se, ainda, uma «preferência» pela integração dos substantivos importados na classe IX, supomos que pela circunstância de que, sendo essa, na opinião que nos parece mais válida, uma classe sem prefixo no singular, se encontra aberta a palavras com os mais diversos inícios. As escolhas de outras classes explica-se normalmente pela coincidência da sílaba inicial da palavra portuguesa com um dos prefixos nominais do Quimbundo que, em consequência, lhe determina a classe. Noutros casos, a anteposição de um prefixo à palavra portuguesa, talvez se possa explicar pela analogia ou pela integração semântica na categoria que a classe engloba. (*ibid.* 89)

Oliveira's explanation is very interesting but requires closer examination. First, if it is true that most loans from Portuguese lose their original number marking, the same does not seem to happen with gender marking, whose marker is apparently interpreted as just another sound in the word. Examples of Kimbundu loanwords from Portuguese provided by Chatelain (1894) seem to support this hypothesis:

| | | |
|---|---|---|
| (3) njanena = 'window' (Chatelain 2001 [1894]:35) | vs. | jinjanena = 'windows' (Chatelain 2001 [1894]:30) |
| (4) diabu = 'devil' or 'demon' (Chatelain 2001 [1894]:36) | vs. | madiabu = 'devils' or 'demons' (Chatelain 2001 [1894]:31) |

In both (3) and (4), what is traditionally called the gender marker in Portuguese[10] is kept[11], while the number (i.e. plural) marker disappears and is replaced by a Kimbundu noun class prefix. So, while Oliveira's explanation (1990:89) seems to account for the addition of these prefixes to Portuguese loans in order to make grammatical (e.g. number) or semantic (e.g. ± animate) distinctions, it does not explain the disappearance of the original number marker, unless we accept that Kimbundu speakers identify and consciously replace the Portuguese

---

[9] Other authors wrote about the influence of Portuguese on Angolan African languages. However, their emphasis was mainly on lexical influence. For examples of such analysis see Martins (1958) and Bal (1979) on the influence of Portuguese on several dialects of Kikongo.

[10] It is not consensual among Portuguese linguists whether gender is actually a morphological category in Portuguese. Whereas traditional grammar (Cunha & Cintra 1995 [1984]; Bechara 2002 [1999]) consider {-a} and {-o} as the morphemes for feminine and masculine, respectively, there is a growing tendency to take them as thematic indexes of a particular noun class (Villalva 2000; Mira Mateus 2003). This issue will be discussed in greater depth in Chapter 4.

[11] Cf. examples like *mulhera* ('mulher'), *favoro* ('favor'), *liquoro* ('liquor'), *papelo* ('papel'), *solo* ('sol') in which {-a} or {-o} are added to the word depending on whether it is feminine or masculine in European Portuguese (Schuchardt 1888:68 - quoted in page 15 - section 2.1.2.).

number markers (suffixes), with their own (prefixes). A phonological explanation seems to better fit the data, i.e. original gender morphemes in Portuguese loanwords are kept due to their vocalic nature ({-o} for masculine and {–a} for feminine), whereas the plural marker (–s) is deleted because all syllables in Kimbundu are open.

In fact, according to Oliveira (*ibid.* 69-89), the choice of the class into which to integrate Portuguese words also seems to be phonologically determined, in that most words are included in the class whose prefix is closer to the first syllable of the Portuguese word. It follows that most words are included in class IX due to its absence of a specific singular prefix, thus allowing for the inclusion of words with many different beginnings. Once again, the examples of Portuguese loans to Kimbundu identified by Chatelain (1894) illustrate Oliveira's point (1990:89):

(5) melekanu = 'American' (adjective) vs. amelikanu or jimelekanu = 'Americans' (Chatelain 1894:256)

Apparently, because the initial <a-> in the Portuguese "americano" matches the plural prefix of Kimbundu class 1 the word is included in that class. However, the singular prefix in class 1 is *mu-* not *me-*, which may help explain why 'American' in Kimbundu has a second plural form starting with the class 9 prefix, i.e. *ji-*.

### 2.1.3. AVP in Portuguese handbooks and grammars

While the first grammars of Portuguese, i.e. Oliveira's (1536) and Barros' (1540), mirror an acute awareness of the social and geographical varieties of Portuguese, they do not cast much light on the type of Portuguese spoken then in Angola. Their statements are too vague and their goal is usually not to evaluate the status and structure of Portuguese there but rather to reinforce some previous statement. For example, Oliveira (1536:45) illustrates the importance of linguistic homogeneity as a source of peace and stability with the following remark about the linguistic situation in the Portuguese colonies:

> ... apliquemos nosso trabalho a nossa língua e gente e ficará com maior eternidade a memória dele e não trabalhemos em língua estrangeira, mas apuremos tanto a nossa com boas doutrinas, que a possamos ensinar a muitas outras gentes e sempre seremos delas louvados e amados porque a semelhança é causa do amor e mais em as línguas. E, ao contrário, vemos em África, Guiné, Brasil e Índia não amarem muito os portugueses que entre eles nascem só pela diferença da língua e os de lá nascidos querem bem aos seus portugueses e chamam-lhes seus porque falam assim como eles.

Oliveira (1536) does not clearly state that the Portuguese spoken in the colonies was different from that spoken in Portugal, but we can infer from his words that the Portuguese in Portugal and the Portuguese born in the colonies made different evaluations of the language they spoke, i.e. the former considered the latter's language to be somehow different, whereas the latter saw no such difference and described their language as Portuguese.

Oliveira (1536) gives no examples of what that "diferença da língua" might consist of but Barros (1540) is more enlightening and refers to specific features of the Portuguese spoken by Africans, namely their pronunciation of some Portuguese words, which he refers to as "barbarismos". These include examples of *aphaeresis* (e.g. "terminar" < determinar), *epenthesis* (e.g. "tôdolos"), *syncope* (e.g. "viço" < vício), *paragoge* (e.g. "guardare" < guarder), *apocope* (e.g. "a mó de falar" < a modo de falar) and *metathesis* (e.g. "apretar" < apertar). These phenomena could also be found in 16th century EP, but, according to Barros (1540), they were much more common in Africa due to language contact, which might help explain why these processes still find parallels today in contemporary AVP:

> E em nenhua parte da terra se comete mais esta figura da pronunciação [i.e. "barbarismo"] que nestes reinos, por causa das muitas nações que trouxemos ao jugo do nosso serviço. (*ibid.* 47)

The incidental description of Portuguese language use in Africa and Angola made by early Portuguese grammarians is not always clear but it does suggest that some linguistic processes resulting from contact occurred in Portuguese Africa as early as 1536 and 1540, i.e. only 52 and 56 years, respectively, after the first Portuguese sailors reached the Congo basin and more than thirty years before the foundation of Luanda in 1576. This makes the almost complete lack of references to these phenomena in contemporary grammars of Portuguese all the more startling.

In fact, while all handbooks and grammars of Portuguese surveyed dedicate at least a chapter to the different varieties of the language spoken around the world (Cuesta & Luz 1971; Teyssier 1989; Cunha & Cintra 1995; Teyssier 1997; Mira Mateus *et al* 2003), a glance through their tables of contents shows that EP and BP are still the only recognized standards of the language. These are the only varieties treated individually (Cuesta & Luz 1971; Teyssier 1997) or compared in separate chapters (Mira Mateus *et al* 2003), the only ones for which an

enumeration of dialects (Cuesta & Luz 1971) and a summary of the debate about its formation (Teyssier 1997) is provided or to which comparative studies are devoted (Teyssier 1989).

All other varieties, including AVP, are merged into generic groups, e.g. "O Português no resto do mundo" (Cuesta & Luz 1971); "O português na África e na Ásia" (Teyssier 1997); "O português de África, da Ásia e da Oceânia" (Cunha & Cintra 1995); "Variação no espaço: o português no mundo" (Mira Mateus *et al* 2003).

Also, whereas great care is given to listing BP's specific features and evolution, not a single line can be found on AVP's specific structural features. In fact, AVP is never referred to as such. The preferred term is Portuguese in Angola, thus diverting attention from any specificity of the language as spoken in that country[12].

Its typological classification is also vague. There is either no reference to it at all (Mira Mateus *et al* 2003) or a distinction is made only between creoles and non-creoles (Cunha & Cintra 1995) and official Portuguese (Teyssier 1997)[13]. The former, are said to be ***languages*** that grew out of a proto-pidgin that developed out of contact between Portuguese and African and Asian languages (Cunha & Cintra 1995:23). The latter, in which AVP is included, are defined as more or less modified ***varieties*** of Portuguese which follow the European standard, except in their oral use (Teyssier 1997).

Part of the problem is that no thought is given to the contact of Portuguese with African languages (Mira Mateus *et al* 2003) or the significance of such contact is disregarded or reduced to the mere borrowing of local lexical items:

> Ao lado de numerosas línguas indígenas fala-se nelas [Angola and Mozambique] um português bastante puro, embora com alguns traços próprios, em geral arcaísmos ou dialectismos lusitanos semelhantes aos que encontramos no Brasil [...]. A influência das línguas negras sobre o português de Angola e Moçambique foi, no entanto, muito leve, podendo dizer-se que abrange somente o léxico local. Pelo contrário, dia a dia, com a incorporação de negros e mestiços na civilização europeia, sente-se aumentar a influência do idioma luso sobre os indígenas. (Cuesta & Luz 1971:144-145)[14]

---

[12] The acknowledgement of the specificities of this variety is made indirectly only, e.g. "E não deixa de ser curioso que por certas particularidades ele [*African Portuguese*] se aproxime do «brasileiro»" (Teyssier 1997:96).

[13] This seems to be a direct legacy from the linguistic works published by Portuguese linguists from the early 20th century until the mid 1970's.

[14] None of the other authors discussed here makes such a blunt statement in this respect, but all of them share Cuesta & Luz's (1971) views (Teyssier 1997; Cunha & Cintra 1995;) even if they allow for a more widespread influence of African languages on Portuguese in Angola, i.e. on the level of morphology and syntax.

Implicit in these words is the homogeneity of the Portuguese language. While apparently recognizing its diversity to some degree, the majority of the authors surveyed here unanimously state that such diversity in no way jeopardises the unity of the language, not even BP, which is agreed to constitute a "specific standard" (Teyssier 1997:91):

> Exceptuando-se o caso especial dos crioulos ... temos, pois, de reconhecer esta verdade: apesar da acidentada história que foi a sua expansão na Europa e, principalmente, fora dela, nos distantes e extensíssimos territórios de outros continentes, a língua portuguesa conseguiu manter até hoje apreciável coesão entre as variedades, por mais afastadas que se encontrem no espaço. (Cunha & Cintra 1995:10)

Concerning the linguistic and sociolinguistic processes underlying the formation of AVP, only Silva Neto (1957 reedited in 1992) departs from the generalized omission or simplification of the aforementioned authors, devoting a considerable amount of thought to issues of language contact, namely to language shift. This, according to Silva Neto (1992), is a highly complex process which involves a relatively long initial stage of bilingualism during which the practical use of both the spoken and written varieties of each of the languages in contact and the social status of the speakers will determine which of the two languages will prevail (*ibid.* 38). According to him this stage of bilingualism determines the structure and typology of the language that replaces other varieties:

> Facilmente se compreende, todavia, que esse contacto, esse período de bilinguismo imprime um cunho na língua que se mantém. As consequências podem atingir uma escala ascendente de graus, que vai até ao esvaziamento da morfologia. Quando sucede este caso extremo, podemos dizer que estamos diante de uma língua mista. (*ibid.* 38)

Silva Neto allows for ascendant degrees of restructuring, the most extreme of which, i.e. those involving the loss of morphological features, are mixed languages. While he does not define the term, he does enumerate the linguistic processes experienced by the prevailing language during the initial stage of bilingualism, i.e. simplification of the grammar, hastening of internal drift due to a weakening of the norm, calques and maintenance of previous substrate linguistic habits[15] (Silva Neto 1992:39).

Another important aspect of Silva Neto's introduction (*ibid.* 42) is his emphasis on the need to distinguish between different types of acquisition of a new language, according to the contexts:

---

[15] In this regard Silva Neto (1992) considers that "a ação do substrato consiste em que os hábitos e tendências lingüísticas de um grupo são mais ou menos perpetuados quando, por via do contacto, há a necessidade de aprender uma nova língua (...) Assim, a acção do substrato não se exercita propriamente na língua, mas no conjunto das tendências a ela imanentes, conferindo-lhes uma direção nova" (*ibid.* 40-1).

> 1 – o caso em que há uma aquisição por indivíduos ou grandes grupos transplantados – quando a assimilação é completa, seja em uma ou duas gerações, seja num período mais longo;
> 2 – o caso em que a aquisição da nova língua é feita por indivíduos que se mantêm no território onde falavam a língua precedente – caso em que verdadeiramente se pode falar de substrato.

Hence, it can be inferred that one must distinguish between the acquisition of Portuguese that led to the emergence of BVP or Portuguese-based creoles (i.e. type 1) and the one leading to the emergence of AVP (i.e. type 2). In fact, the emphasis on the study of the sociological contexts inherent to each language contact situation is the most distinctive feature of Silva Neto's work, which is summarized in the quote below:

> Tudo dependerá, pois, de certas condições sociológicas, tais como o ambiente social, o contacto e a interação, o maior ou menor número de não aloglotas, o isolamento, o maior ou menor desejo de ascensão social. (*ibid.* 45)

In fact, when discussing the diffusion of the Portuguese language in Africa, and especially in Angola, Silva Neto gives great thought to the influence of sociolinguistic factors in the formation of this variety, even if not stating it explicitly. For example, following the Portuguese historian Oliveira Martins[16], Silva Neto (1992:430-1) states that different types of colonial settlement imply different means of transmission of the European language and therefore different linguistic outcomes:

> Nas feitorias, o contacto é ainda espaçado e vacilante: propicia a formação de uma *lingua franca* intermediária. Nas fazendas, o contacto é íntimo e decisivo: proporciona a formação do *crioulo* como instrumento único de comunicação. Nas colónias, a situação é bem mais complexa, pois nelas se estabelece uma camada de *élite* branca.
>
> Essa camada procura, ou não, manter e desenvolver os padrões culturais e os valores da mãe-pátria.
>
> O destino linguístico de uma terra conquistada não dependeu, absolutamente, do substrato, mas sim do modo de transmissão da língua europeia.
>
> Esse modo de transmitir o idioma é que lhe vai definir o futuro carácter. Se apenas se estabelecem relações comerciais ... só há o interesse de mútua compreensão, - desenvolve-se um falar de tipo crioulo.
>
> Mas, se pelo contrário, se busca plantar uma nova pátria; se há fusão sob o signo da cultura dos mais bem dotados; se à transmissão oral da língua se acompanha a transmissão através da escola, - então mantém-se perduravelmente o essencial da língua européia, que apenas oferecerá matizes de variantes. [...] Esses fatos repetiram-se no Brasil e na Hispano-América.

[16] While Oliveira Martins is the only source referred to by Neto (1992:430-1) a similar classification of creoles according to the circumstances of their formation can be found in Reinecke (1937), and later on in Chaudenson (1979) and Bickerton (1986). For a detailed analysis of the classifications proposed by these authors see Holm (2000:40-41).

Taking into account that Angola had different colonial patterns of settlement at different moments of its history (i.e. *feitorias* from the 15th to the 17th century and *colónia* from the late 19th century onwards), one can, according to this line of thought, pose two hypotheses to account for the origin of AVP: (1) it is the result of a *lingua franca*, i.e. a "falar de tipo crioulo", or (2) it is a mere regional variety of Portuguese. Accepting the first option as true, one could state that AVP's synchronic structure sprung from the "português dos negros" or "falar crioulizante"[17], which later on developed into a "semicreole", i.e. a language variety that at a certain stage of its history underwent some degree of creolization only to later on decreolize as a consequence of increased contact with the standard:

> Nos crioulos há vários graus de aprendizagem, pois, segundo as circunstâncias, o primitivo falar xacoco mantém-se ou é aos poucos renovado pelo sangue novo da língua européia. De geração em geração, graças sobretudo à escola, vai-se aperfeiçoando e enriquecendo a primitiva fala de emergência [...] Daí o admitir-se a existência do *semicrioulo*, ou seja, o estágio mais aperfeiçoado da primitiva aprendizagem. Ele exemplifica-nos o choque entre o falar europeu e o crioulo. Este vai sendo, pouco a pouco, invadido por palavras e giros do falar das pessoas socialmente mais bem dotadas. O semi-crioulo encerra, pois, formas e torneios semicultos. (Silva Neto 1992:437)

However, none of the descriptions of Portuguese spoken in Angola quoted by Silva Neto seems to validate this hypothesis. In fact, in his description of the diffusion of the Portuguese language in Africa, Silva Neto (1992) does provide a large body of testimonies reflecting Portuguese and foreign travelers' general perception of the language as spoken by Africans in Angola, but the earliest impressions quoted by Silva Neto are in straight opposition to the understanding of AVP as resulting from decreolization (which is implied by his definition of "semi-creole" above), because all of them range from describing Africans as quite knowledgeable about Portuguese (*ibid.* 513) or as least as possessing a reasonable knowledge of the language (*ibid.* 516). These reports are all the more relevant as they date to the 16th and 17th centuries, when the Portuguese presence in Angola was guaranteed by means of the *feitorias*, which according to Silva Neto (*ibid.* 430-1) should have resulted in the creation of an intermediary lingua franca. Only in the late 17th, 18th and early 19th century is the Portuguese spoken by Angolans described as corrupted Portuguese (Silva Neto 1992:517).

---

[17] Silva Neto (1992:436) defines the "português dos negros" or "falar crioulizante" in the following terms:

> Este não constitui um conjunto de caracteres definidos. É meramente individual: exemplifica-nos a fase inicial, e o primeiro contacto, que pode preparar o advento do crioulo [...] consiste na algaraviada ocasional das tribos que, resistindo à assimilação, mantêm intactos os seus padrões culturais, inclusive, é claro a própria língua.

> La langue portugaise corrompue s'y est conservée jusq'à présent, produit un jargon ou langue franque que presque tout le peuple entend, parle, de sort que ceux que savent le portugais, n'ont pas besoin d'interprète. (Labat 1730, quoted by Silva Neto 1992:515)[18]

Later on, Silva Neto (1992:517) quotes Corrêa, an historian of Angola, who referring to 18th century Angola describes it as mainly Kimbundu speaking:

> As senhoras, custumadas a fazerem-se entender às suas escravas por esta linguagem, são verbosas nas conversações familiares e mudas nas públicas assembleias. A ..., baronesa de Mossâmedes, ardendo nos desejos de fazer brilhante o tempo do governo de seu ... esposo, as convocou, e reduziu a aparecerem vestidas ao uso da Europa, atraindo primeiramente à sua presença algumas meninas das principais famílias, as quais educou debaixo dos preceitos e maneiras europeanas, mandando-as ensinar a costurar, a bordar, a ler, a escrever, a contar, a música, a dançar e, por consequência, a falar... Os homens falam português, e são elegantes no ambundo. (Corrêa quoted in Silva Neto 1992:517)

Regarding the 19th century, Silva Neto states that there is good information about the pronunciation of Portuguese in Angola and quotes a grammar of Kimbundo:

> A maior frequência da lingua angolense entre os naturaes do paíz tem modificado a pronunciação da portugueza que é falada en Ngola (Angola) mais ou menos correctamente pelas pessoas civilisadas com a branda pronúncia do seu idioma patrio, mas com a acentuação das terminações similhantes à de alguns paulistas[19] (no Brasil), não tão forte, mas parecendo apenas que as vozes que terminam as palavras tem o som prolongado, como se fossem dobradas ou tivessem éco. Este vicio na pronuncia da lingua portugueza é devido à frequência das exclamações ou às partículas de interjeição da lingua Nbundu, sempre empregadas no fim das palavras ou das orações. (Oliveira & Francina quoted by Silva Neto 1992:517)

It is clear from the quotes above that most Angolans spoke an African language rather than Portuguese and that when they did speak Portuguese they used it fairly well, albeit with phonetic interference from their mother tongue. What is described above is not a creole undergoing decreolization but rather a whole new variety in the making.

While not developing a full theoretical model to account for AVP's origins and synchronic structure, Silva Neto (1992) nonetheless provides some valuable insights into the cause/effect relationship between sociolinguistic factors and different degrees of restructuring, even if not clearly defining them. Furthermore, he quotes a significant number of sources describing the

---

[18] While the words above constitute an invaluable source of information about 18th century Portuguese in Angola, they should perhaps be analyzed with caution, especially because these testimonies are usually made by non-native speakers of Portuguese whose knowledge of the language is seldom known and who some times report hear-say rather than their own personal experiences.

[19] Cf. with Schuchardt's reference to the "return of vast number of emancipated slaves from Brazil to these regions" (1888:62-3) quoted on page 16 of this thesis.

linguistic situation in Angola from the 16th to the 19th centuries as well as the language policies enforced there.

### 2.1.4. Works on Angola's Bantu languages

The first linguistic research on African languages was undertaken in Angola as early as the seventeenth century (Bonvini 1994:127) and intensified during the eighteenth century, in part due to the increment of missionary action in Angola (especially by Jesuits and Capuchins) and the consequent growing contact with Angolans' native languages. Since the majority of these languages were completely unknown to missionaries, grammars and dictionaries were published to help them learn these languages (Cole 1971:4). These publications were essentially descriptive in nature and contained little or no information regarding the use of Portuguese in Angola.

While nineteenth-century grammars and dictionaries are also not much help in describing AVP, they nonetheless offer an important insight into the sociolinguistic scene in the country at the time of their publication by highlighting the widespread use of Angolan national languages as lingua francas in the territories administered by the Portuguese. This is shown by most of their titles, e.g. Cannecatim (1804) and Batalha (1891) and Nascimento (1893), and stated by the authors themselves. Cannecatim (1804:iv-v) asserts that due to Kimbundu's status as a lingua franca, its use by missionaries and priests, administrators and traders was crucial to the success of the Portuguese interests in Angola because this allowed them to avoid dependence on interpreters:

> ... a intelligencia da Lingua Bunda, ou geral do Reino de Angola, he utilissima, e necessaria aos Ecclesiasticos no exercício do seu ministerio; aos Governadores, e Magistrados na Regencia do Estado, e Administração da Justiça; aos Chefes Militares no acerto do seu Commando e na felicidade de suas operações; aos Commerciantes em fim no manejo do seu negocio, sendo huma ruina, e huma desgraça, que todas estas pessoas não vejão o objecto de suas funções, senão ao travéz da opaca sombra de um Negro interprete. (Cannecatim 1804:iv-v)

This is further reinforced by Nascimento (1893:xii), who claims that "quem souber a língua dos bienos [i.e. *Umbundo*], pode viajar por toda a Africa austral com a certeza de ser entendido por todos os povos ba-ntu"[20]. Hence he dedicates his work to the "negociantes, viajantes, funccionarios publicos, missionarios, militares, etc., que tenham de viver em contacto com os indigenas" (Nascimento 1893:xii). This shows that the more the Portuguese advanced into the

[20] Umbundu is to this day one of the most important *lingua francas* in Africa (Moseley & Asher 1994:289-309).

hinterland, the less their language was of any use in contact with the natives. As Chatelain (1894:viii) puts it: Portuguese was to the majority of Angolans "what Latin is to the Lusitanian peasant".

The predominance of the local languages over Portuguese as lingua francas continued well into the mid-twentieth century. However, due to the increased influx of Portuguese settlers into the interior, the growing number of "assimilados" and the consequent expansion of the area of influence of Portuguese, the goal of most grammars and dictionaries of Angolan native languages was now not only to enable Europeans and assimilated Angolans to communicate with the natives, but above all to spread the Portuguese language among the latter (Maia 1994 [1961]:viii). A practical knowledge of the native languages was perceived as a crucial means to achieving linguistic and cultural assimilation of Africans (Maia 1994 [1961]:vii; Maia 1957:viii), for despite Portuguese efforts the linguistic context described by Cannecatim (1804) remained accurate:

> ... os primeiros missionários têm-se visto a braços com a dificuldade de se compreenderem ou de se fazerem compreender do povo nos pontos em que a Língua Portuguesa não atingiu ainda o necessário incremento. Aos funcionários Administrativos sucede por vezes o mesmo. Os srs. Aspirantes e Chefes de Posto veêm-se coagidos, pela necessidade de resolução de casos indígenas, ao estudo prévio das línguas nativas para não recorrerem sempre à intervenção de cabos de cipaios que lhes servem de valiosos intérpretes das questões diárias. (Ex-seminar student in the introduction to Maia 1994 [1961]:ix)

Curiously enough, none of the works surveyed here makes reference to contact-induced changes in Portuguese, but rather to significant change and dialectal differentiation within the African languages. For example, Chatelain (1888-89; 1894) clearly distinguishes between three dialects of Kimbundu based on the ethnicity of their speakers and their degree of contact with the Portuguese: the "intermediary dialects", which he defines as mixtures of Kimbundu with Kikongo or Umbundu (*ibid.* xiii) spoken by "tribus independentes e até aqui refractarias à civilisação e ao jugo portuguez" (*ibid.* xii) and the two dialects of what he terms "kimbundu proper" (*ibid.* xiii), i.e. the Kimbundo spoken in the hinterland of Luanda and the Kimbundu spoken in the city of Luanda. He claims that the former is the purest and the variety closer to the original language of the Ndongo kingdom, as opposed to the latter, "the form of speech in daily use among Loanda natives, needlessly mixed with Portuguese elements" which "offers poor material for the study of the genuine Ki-mbundu" (Chatelain 1894:v).

Tavares (1915) further elaborates on this topic and while he tries to minimize the contact-induced changes undergone by the Angolan Bantu languages, the phenomena he refers to seem to indicate significant restructuring, i.e. the tendency to use abbreviated forms of agreement, the

acquisition of whole new sets of grammatical words and the substitution of native words by the assimilated forms of their Portuguese equivalents:

> A língua [i.e. *Kimbundu*], principalmente nos centros mais populosos e cultos, além de se apresentar com acentuada tendência para a realização de formas contractas, tem modificado, pôsto que de leve, algumas das suas leis de concordância suprimindo certos elementos morfológicos e adquirindo outros. [...] O contacto de quatrocentos anos com a língua portuguesa ... não alterou a pureza da língua. Os elementos estranhos, que são na verdade poucos, tem sido adaptados à gramática nativa: termos para os artigos que antes de nós não conheciam, alguns verbos, advérbios, preposições e conjunções, sendo porêm de notar que foram desprezados alguns antigos termos nacionais, os quais foram substituídos pelos nossos, inteiramente assimilados. (Tavares 1915:vii-viii)

The reference to contact varieties of local Angolan languages as a result of both contacts with Portuguese and with other African languages is recurrent on more contemporary sources on language use in Angola, despite the fact that specific linguistic traits are seldom provided. As far as the first type of contact is concerned, Oliveira (1990:74) refers to the existence of tertiary creoles[21], such as Olumbali[22]. The same author also refers to several "mixed languages", such as the Quimbundo de Nambuangongo and Bângala[23].

To the best knowledge of the present author, no grammars or dictionaries of Angolan Bantu languages were published from the 1960s to Angola's independence from Portugal (i.e. 1975). After independence, Angola adopted Portuguese as the official language, but the great linguistic diversity of the country and the general population's limited knowledge of Portuguese led Angolan rulers to enforce measures with view to the description of the indigenous languages and their subsequent use in literacy and schooling, e.g. the establishment of the *Instituto Nacional de Línguas* in 1979. Despite the fact that these measures had few practical results[24], they nonetheless led to a growing debate about the linguistic situation in the country. In the

---

[21] The term was first used in Valkhoff (1966), where it was defined as the result of "the Portuguese influence on Bantu languages" (p. 34). This is the meaning in which the term is used here.

[22] This is a language created in Moçâmedes (present day province of Namibe) by the "quimbares", i.e. "todos os que surgem associados a europeus, designadamente pelo trabalho, deixando-se influenciar pela sua cultura" (Oliveira 1990:74).

[23] The only information provided by Oliveira (1990:79-80) about these "mixed languages" is transcribed below:

> Uma língua mista atestada por Atkins, na fronteira Norte com o Quicongo é o Quimbundo de Nambuangongo, variedade «peculiar a esta região. A estrutura gramatical parece-se mais com a do Quimbundo do que com a do Quicongo, mas a inversa é verdadeira para a fonologia. Estamos em face de uma língua mista que revela um alto grau de interferência recíproca das duas línguas também no vocabulário. Língua mista é também considerado por Atkins o Bângala, resultado da interferência intensa entre o Quimbundo e o Quioco, e que se situa na fronteira entre estas duas línguas.

[24] My fieldwork in the country made it clear to me that the practical goals of these measures have yet to be achieved, i.e. there are no specialized teachers of national languages; there are very few, if any, text books to teach them; the orthography of the national languages remains an unsettled issue, etc.

period immediately following independence and during the 1980s, the focus of debate was essentially on Angola's African languages, termed "national languages", as opposed to Portuguese, i.e. the official language. The most important contribution of this stage of the debate was the identification of the most representative national languages[25], i.e. Kikongo, Kimbundu, Cokwe, Umbundu, Mbunda and Kwanyama, and the proposal of alphabets for each of them (Instituto Nacional de Línguas 1980). By 2000 the awareness that the number of speakers of Portuguese in Angola, both first and second-language speakers, had increased brought the debate about the linguistic situation in Angola to a stage in which Portuguese was included among Angola's national languages[26]:

> Temos que reconhecer que ainda não está suficientemente claro para todos nós, Angolanos, o que será uma língua angolana. Em nossa opinião, esse termo aplica-se a todo o meio de comunicação, usado no país, de origem local ou não local mas que, como acontece com a língua portuguesa, ganhou esse direito, pelo seu percurso histórico no nosso país. (Mingas 2002)

> As áreas de difusão da língua oficial, a Língua Portuguesa, correspondem em certa medida, aos limites fronteiriços do País. Ela cobre toda extensão do território Nacional e é falada de Kabinda ao Cunene, do Mar ao Leste. Implantada em Angola há mais de cinco séculos, enraizou-se profundamente na sociedade. Hoje, para além do seu estatuto de língua oficial, ela é a língua materna de muitos angolanos e com o seu alastramento constitui a língua nacional no sentido pleno e veicular para todos os angolanos, embora o grau de domínio não seja igual para todos. (Fernandes & Ntondo 2002:19).

More important than the acknowledgement of Portuguese as one of Angola's national languages is the underlying assumption that due to contact with African languages, a whole new restructured variety of Portuguese has been forged and has spread throughout Angola, i.e. AVP:

> A coabitação com as demais línguas angolanas originou o surgimento de uma variante nacional do português ao que chamaríamos *"português vernacularizado"* que, massificando-se, tornou-se *veicular*. (Fernandes & Ntondo 2002:19)

> ... a língua portuguesa realiza-se numa situação de contacto de línguas onde a mesma coexiste com as demais línguas Nacionais, não bantu e bantu. Hoje fala-se mesmo do português angolano por ter sido apropriado pelos angolanos e um bom número dentre eles, sobretudo citadinos têm-na como língua materna. (Fernandes & Ntondo 2002:103)

Hence, according to Fernandes & Ntondo (2002:19) it is AVP and not some transplanted standard of EP which is a national language in Angola and which is either the first or the second

[25] The term 'national language' referred to the indigenous languages as opposed to Portuguese.

[26] For a more detailed analysis of this ongoing debate in Angola see Cabral (2005:20-3 and 26-8).

language of Angolans, even if the authors do associate AVP with a specific ethnolinguistic group, i.e. *descendentes*[27].

Fernandes & Ntondo (2002) do not clarify whether the Portuguese they describe is EP or AVP, but the fact is that the pronunciation suggested by their spelling for some Portuguese words does not match that of EP. For example, for the name of Portuguese letter 'b' the authors suggest the pronunciation [bé], with the low mid front vowel [ɛ] instead of the high mid front vowel [e] in EP (2002:106). The pronunciation suggested for the <a> in "anel" is [a] rather than [ɐ] (2002:111) in EP. No references are made to morphosyntax.

Despite this, the relevant fact is that AVP is not some theoretical construct made up by foreign linguists, but rather the product of Angolans' own perception of the variety of Portuguese spoken in Angola as distinct from EP, even if its description remains to be done.

### 2.1.5. Specific works on Angolan Vernacular Portuguese

As stated in section 2.1.2., specific work on the topic appeared only after the country's independence from Portugal in 1975 and it was essentially concerned with the description of lexical borrowing and transfer. The first references to specific features of AVP other than lexical appeared only in the 1980s in literary criticism of the work of politically engaged Angolan writers who had taken an active role in the country's fight for independence during the 1960s and 1970s by mirroring in their literature the socio-cultural context of their country (Menezes-Leroy 1989:102). Language played a crucial role in this. While some only wrote in standard EP (e.g. Manuel Rui, Pepetela, Rui Duarte de Carvalho), others, like Luandino Vieira and Uanhenga Xitu, tried to echo the language of the common people, or what Óscar Ribas called "popular Portuguese" (in Laban 1991:33-4), whose features included, for example, use of Kimbundu verbal roots with Portuguese verbal endings, local phonetic variants of Portuguese words, lack of subject-verb agreement, lack of number and gender agreement.

To some extent this literary language echoes tendencies in the oral speech of most Angolans, but caution is needed in taking it at face value, since it illustrates the contact of Portuguese with Kimbundu only. Hence, the phenomena transferred from African languages to Portuguese which they illustrate may be similar to that of "*la langue populaire de la region de Luanda*"

[27] The group is composed of the descendants of Europeans (mainly Portuguese) and the descendants of Europeans with Africans. According to the authors, the former speak Portuguese as their mother tongue, whereas the latter can be divided into those who speak Portuguese as their first language and those whose mother tongue is a Bantu language (Fernandes & Ntondo 2002:101-2).

(Laban 1982:61), i.e. the "*linguagem dos musseques*" (Laban 1991:31; Perl & Hundt 1991:51), where Kimbundu is actually the predominant language, but not necessarily similar to AVP as a whole (Perl 1989a:14). Moreover, the authors themselves were usually neither native speakers of Kimbundu nor of the variety of Portuguese they tried to mirror:

> Embora eu conheça mais ou menos o português popular, mas depois tenho de estar a acertar: «Vamos cá ver se é assim... Não é bem assim!» e lá corrijo uma palavra... «Mas não é assim...», Bom a não ser que seja uma criatura que fale bem o português popular, que não é o meu caso. E então, depois de escrever o conto em português, a pessoa, o informante ouvia-o ler para ver se era mais ou menos assim, assim – uma frase ou outra: «Não é bem assim...» E eu corrigia. Quase todos os contos foram feitos assim, para lhes conceder maior cunho de autenticidade. (Óscar Ribas in Laban 1991:34)

In short, the "português angolanizado" (Macedo 2002) used by the Angolan writers of the 1960s and 1970s was not intended to mirror the exact speech of the common people, but to empower and dignify all those who in Angola used non-standard varieties of Portuguese and who had so far been discriminated against in education and jobs for doing so (Macedo 2002).

Indeed, it was this concern over the educational consequences of speaking a "português angolanizado" rather than standard EP, i.e. the official language, that led to the emergence of the first strictly linguistic studies on AVP in the 1980s and 1990s. These were essentially about the teaching of Portuguese as a second language in Angola (Marques 1990 [1983]; Cuesta 1990; Ançã 1998) and included interesting insights regarding the linguistic situation in the country. First, they showed that less than 20% of the country's population spoke Portuguese natively, i.e. essentially residents in the coastal cities (Cuesta 1990:15), against the 85% (Kounta 1980:1) or 80% (Cuesta 1990:15), especially in the rural areas, who spoke one or more Bantu languages natively and had little or no knowledge of Portuguese[28]. Second, these studies provided examples of linguistic interference in the speech of non-native speakers of Portuguese, offering as a possible explanation the contact between Portuguese and the Bantu languages:

> A realização da língua portuguesa no nosso país [i.e. Angola] dá-se numa situação de plurilinguismo ... isto significa que o fenómeno do contacto linguístico e o que dele decorre – como por exemplo, as interferências linguísticas – não pode ser ignorado nem negligenciado. (Marques 1990 [1983]:209)

[28] This fact has been acknowledged by many since colonial times, e.g. in 1968 Barreto presented a report about the teaching of Portuguese as a second language in Africa at the *I Simpósio sobre a Língua Portuguesa Contemporânea,* in which he acknowledged that 70% of the total population in Guinea Bissau, Angola and Mozambique had little or no knowledge of Portuguese (Barreto 1968:147). Of course, the lack of efficiency of second language teaching of Portuguese in Africa in colonial times was seen as a problem because it prevented assimilation, whereas from the 1980s onwards it has been a matter of concern because it prevents the majority of the population from attending school.

According to Marques (1990 [1983]), this is illustrated by the lack of number and gender agreement in the noun phrase in the speech of Angolan speakers with little knowledge of Portuguese, since number and gender in Bantu are expressed by adding prefixes to the elements of the noun phrase, Angolans interpret the Portuguese article as the equivalent of those prefixes and therefore do not mark the plural or feminine at the end of nouns to establish agreement with its determiners in instances like "os pai" and "as camarada" (Marques 1990 [1983]:219).

Interestingly, in Marques (1990 [1983]) the examples of interference provided, e.g. the lack of number and gender agreement in the noun phrase[29], are not taken as "errors" but as specific features of what she calls "língua portuguesa falada em Angola" (1990:214) or "português angolano" (*ibid.* 215), i.e. AVP. In fact, she actually suggests a plan for its description with view to the establishment of a standard and the consequent revision and restructuring of descriptive and normative grammars of Portuguese:

> – Um levantamento do português fundamental (mais usual) em Angola e o que ele traduz em termos de uso pela população angolana (estatísticas por áreas geográficas):
>  a) compreensão;
>  b) compreensão e expressão oral
>  c) compreensão, expressão oral e expressão escrita.
> – A descrição do português de Angola (características fonéticas, morfo-sintácticas e semânticas).
>  Esta análise descritiva, a longo prazo, vai permitir igualmente delimitar:
>  a) as marcas do português de Angola que decorrem da evolução natural da língua (o dinamismo e evolução das línguas vivas);
>  b) as marcas do português de Angola que decorrem do seu contacto com as línguas nacionais (interferências, etc).
>
> (*ibid.* 215)

There are two very interesting implications in Marques' (1990:215) words. First, in acknowledging that AVP includes dialectal differentiation and that its speakers have different degrees of proficiency, she implicitly emphasizes that AVP is not a unitary language variety and consequently that any serious attempt at describing it must take this fact into consideration. Second, any theoretical attempt at classifying AVP must pay equal attention to both contact with Bantu languages and the natural drift of Portuguese.

While Marques' (1990 [1983]) plan has still not been implemented, it did propose a path for future investigations. Ançã (1998) is a case in point, in that it gave an interesting contribution to the better understanding of Angolans' linguistic competence in Portuguese and to the identification of their main difficulties in using the language. With these goals in mind, she gave

[29] The Angolan linguist also identifies other particularities of the "português de Angola" (1990 [1983]:215), namely lack of subject/verb agreement. For other features cf. Marques 1990 [1983]:217-23.

a test to 64 Angolan and Cape Verdean students[30] (*ibid.* 1032-3) in which they were expected to recognize and produce correct Portuguese constructions. While most Cape Verdean students claimed Cape Verdean Creole was their mother tongue, the case with Angolans was more complex: 38 students stated Portuguese was their mother tongue and 26 replied theirs was either Kimbundu, Kikongo, Fiote or Umbundu. One would expect that all non-native speakers of Portuguese (both Cape Verdean and Angolan) would experience difficulties in most questions, as opposed to Angolan native speakers of Portuguese. However, the results showed no such distinction, for both groups (both native and non-native) experienced more difficulties on exactly the same questions: those related to the plural form of nouns, uniform adjectives[31], gender inflections in adjectives, word formation and identifying ungrammatical sentences. This is very interesting and completely in line with my observations of Angola's linguistic reality during my fieldwork there:

> Embora, aparentemente, as línguas banto estejam a perder vitalidade na capital, elas continuam presentes no português que hoje se fala em Luanda, tanto a nível lexical, como morfo-sintáctico e fonético. (Ançã 1998:1035)

Ançã (1998) concludes that partial knowledge of Portuguese is particularly visible in syntax, but that not all difficulties are of the same sort. Hence, she considers that most of the grammatical difficulties discussed above are what she calls "interlinguistic difficulties", i.e. they result from language contact and interference, whereas the difficulties in identifying ungrammatical sentences are what she calls "intralinguistic difficulties", i.e. resulting from hypercorrection, overgeneralization and analogy (*ibid.* 1039).

The first truly linguistic study of Portuguese language use in Angola was also published in the 1980s, i.e. Mendes (1985). The goals of the study were threefold:

> ...verificar quais as características actuais da língua portuguesa falada na situação de plurilinguismo da comunidade de discurso angolano, analisar as mudanças nela operadas em relação à norma e reflectir sobre as suas causas. (Mendes 1985:1)

Hence, Mendes (1985) starts with a detailed description of the sociolinguistic situation in Angola in the 1980s, which in her opinion was highly complex due to the coexistence of three

[30] Only students living in the capitals of these countries took the test, Luanda and Praia, respectively. All of them were studying at "Instituições de Formação de Professores".

[31] In Portuguese traditional grammar this refers to those adjectives that possess a single form to both the feminine and the masculine, e.g. *feliz* ('happy'); *jovem* ('young'), etc.

different linguistic groups, i.e. the national languages[32], Portuguese, and other foreign languages:

> Uma grande parte da população é unilingue ou em português (centros urbanos) ou numa língua nacional (zonas rurais, principalmente); outra parte é bilingue e alguns são até plurilingues, empregando-se: uma língua nacional mais português; duas ou várias línguas nacionais; português mais outro (ou outros) sistema(s) linguísticos(s) estrangeiro(s); língua nacional mais língua portuguesa mais outra língua estrangeira. (*ibid.* 34)

The group of monolingual speakers includes Angolans (monolingual in a national language, mainly in the rural areas) and Portuguese cooperators and Angolans whose parents are Portuguese (monolingual in Portuguese, mainly in urban areas). The group of bilingual and plurilingual speakers includes Angolans, whose one of parents is Angolan and the other foreign, Angolans who studied abroad, and foreigners in general (*ibid.* 32).

While this characterization may be a bit of a generalization, Mendes' (1985) prediction of different sociolinguistic outcomes for the rural centers and the urban centers in the long run shows great linguistic awareness. In the rural areas, due to the rarity of interlinguistic contacts, the speakers' resistance to the social pressure of Portuguese, the inaccessibility of some regions and their distance from the capital, Mendes (*ibid.* 46) predicted that monolingualism or multilingualism in national languages would prevail to the detriment of bilingualism in Portuguese and Bantu languages. In the cities, the higher social, economic and cultural prestige of Portuguese, the growing concentration of services and refugees, the exclusive use of Portuguese in schooling and the war would in time lead to monolingualism in Portuguese through language shift (*ibid.* 40-3):

> ... a situação linguística actual em Angola tenderá a longo prazo para uma mudança de línguas, ou seja, que algumas das línguas nacionais angolanas desapareçam, evoluindo-se para uma situação de unilinguismo em proveito da língua portuguesa (naturalmente já distanciada da actual norma portuguesa). (Mendes 1985:46)

As far as the development of AVP is concerned, Mendes (1985) advocates that "O português falado em Angola parece ter seguido, na sua formação, dois caminhos: um conservador, que se tem desenvolvido lentamente, e outro inovador, a que condições sociopolíticas imprimem, por vezes maior velocidade" (1985:184).

The factors accounting for AVP's conservative character are essentially (a) the conservative tendency of transplanted languages, (b) dialect leveling in consequence of the different geographic backgrounds of the Portuguese settlers and their concentration in small centers, (c)

[32] The term here is synonymous with indigenous language.

the archaic Portuguese spoken by the majority of the settlers, mostly from rural areas in Portugal with little or no contact with cultural centers, and (d) the role of missionaries and schools in spreading the literary norm of EP. The different word order of object personal pronouns in AVP, the use of the preposition *em* with verbs expressing movement, and the pronunciation of final "e" as [i] are some features used by Mendes (1985) to illustrate AVP's conservative character.

AVP's innovative character depends on the acceptance by the community of individual innovation and on the role of several different social, political and historical factors which can accelerate or retard that acceptance. Mendes (1985) identifies three favorable stages to linguistic innovation in Angola: "a primeira logo no início da colonização, a segunda no último período colonial e finalmente a terceira é a fase actual, após a Independência" (*ibid.* 196). The first and the last stages are the most relevant. The first because the contact with a new reality resulted in lexical borrowing from Bantu languages, disappearance of certain words that were no longer needed, semantic and grammatical transformation of the most often used words, and imperfect learning of Portuguese by Angolans. The last stage was highly relevant because it brought about the mass movement of people due to the war and political, technical, social, economic and cultural changes favoring innovation[33], especially in the capital.

Thus, according to Mendes (1985), AVP appears to be the product of a process similar to the one that originated BVP. In fact, when comparing both varieties, she states that they share several structural features, i.e. modification of EP phonology, lexical borrowings and calques, and the intensification of linguistic trends prior to the transplantation of the language. Still, she stresses that due to different sociolinguistic backgrounds "o português no Brasil viria a sofrer rumo diferente do que seguia em Angola" (*ibid.* 201).

According to Mendes (1985), *interlinguistic interference* is the one linguistic process which accounts for the growing differentiation between AVP and standard EP:

> Sabemos que dada a situação sociolinguística actual de Angola as interferências tendem a aumentar provocando divergências na língua portuguesa falada nesse país e afastando-a progressivamente do português-padrão e que este processo de evolução continuará por certo tempo ainda mas pensamos que neste momento já pode ser feito um estudo sincrónico dessas interferências. (Mendes 1985:66)

Lexical borrowings from Bantu (especially Kimbundu), BP and other foreign languages; lexical and semantic innovations in Angolan Portuguese and morphosyntactic and phonologic Bantu

[33] Mendes (1985:198) includes in these changes the transformation of Angolan social structure, bigger freedom of speech, more contact of Angolans with the outside world and improvements in education.

transfer into Angolan Portuguese are the main consequences of this process of *interlinguistic interference*, for which Mendes provides an extensive and valuable list of examples (*ibid.* 68-159). Extremely important and exhaustive as this is, Mendes's (1985) enumeration of transfer phenomena needs analysis, as all her examples are taken from a literary corpus[34]. Other shortcomings include the use of examples that can be misleading in that they are not exclusive of AVP[35], and the use of vague and not very technical terms to refer to the phenomena she surveys, be it lexical and semantic innovations in AVP (1985:102-31)[36] or morphosyntactic interferences in AVP (*ibid.* 133-57), i.e. "deslocamentos"[37], "acrescentamentos"[38] and "dificuldades"[39], which is not complemented with individual analysis or explanation of the examples.

As far as lack of agreement is concerned, Mendes (1985:149-51) identifies five instances:

(a) lack of number and gender agreement between the determiner and the noun:

(6) "Porque não tira o fotografia."
(7) "Tem muitas mulher_."

(b) lack of number and gender agreement between the noun and the adjective:

(8) "As coisas estão muito fresco_"
(9) "abriu a cancela pequeno do fundo do quintal"

(c) lack of subject/verb agreement:

(10) "as pessoas paga_ renda no fim do mês"
(11) "tu quer_ ir mesmo com aquele rapaz"

---

[34] See the discussion above in this section about the shortcomings of using literary works as linguistic corpora.

[35] For example, Mendes (1985:110) states that *afrente*, *agente* and *derepente* are examples of agglutination of the Portuguese expressions *à frente* (i.e. 'ahead'), *a gente* (i.e. 'we'), *de repente* (i.e. 'suddenly'), which is true, but the examples she provides seem to indicate that these are not new Angolan Portuguese words but rather phonetic agglutinations identical to ones found in informal EP.

[36] For example, what does "deformação" and "aportuguesamento" mean? Have "alambamento" (i.e. from Kimbundu "kilembu" meaning "dowery") and "funjada" (i.e. from Kimbundu "funji", referring to a large quantity of "funge", an Angolan typical dish) undergone the same morphological processes?

[37] This refers essentially to the different use of the personal pronouns in Angolan Portuguese, e.g. "já vou-me embora" (Mendes 1985:146); the use of the direct object before the verb, e.g. "mas cinco minutos nem que passaram" (*ibid.* 147); non-inversion of the subject in interrogative sentences, e.g. "Onde ele está?" (*ibid.* 148).

[38] This refers essentially to the insertion of words in the sentences for emphasis.

[39] The meaning of the word is unclear, but under this heading she includes difficulties in using relative clauses and sentences (e.g. "o nosso esquadrão **que** o Comandante lhe chamamos Kwenha", *ibid.* 155); comparative forms (e.g. "É **mais pior** de ver" , "levantou a cabeça para ver **mais bem**", *ibid.* 155); object pronoun with preposition (e.g. "Teu patrão um dia correu **na minha trás**", *ibid.* 156); certain verb tenses and modes like the imperative (e.g. "Psiu, **vens** cá", *ibid.* 156) and the subjunctive (e.g. "dar nos lombos de todos que **fazerem** barulho", *ibid.* 157); and of passive sentences (e.g. "impedir que os garotos mais fracos **fossem batidos** pelos mais fortes", *ibid.* 157).

(12) "eu quando veio"

(d) lack of tense agreement

(13) " Tanto estão bem connosco de manhã, como à tarde estarem a desejar a morte"
(14) "O senhor Felito compra medicamentos e depois venha com o enfermeiro"

(e) use of different verb persons to refer to the same interlocutor.

(15) "O senhor quer ficar bem com a tua mulher"
(16) "Querem-te roubar o ovo da sua mãe"

The reduced number of examples and the lack of individual analysis and explanation of each of them is unhelpful, especially when the only explanation provided is reduced to the interference of Bantu languages and simplification of the Portuguese grammar in the terms that follow:

> No quimbundo o número recebe tantas variações quantas as classes (são dez as classes de substantivos e há dez prefixos que lhes correspondem). Por darem menos importância ao género aparecem no seu discurso discordâncias como "no mesmo barriga". (Mendes 1985:214).

In spite of the shortcomings mentioned above, in which the implicit belief in the homogeneity of AVP (Mendes 1985:200) must be included, Mendes (1985) provides a remarkably lucid appreciation of the Angolan linguistic reality:

> Não sabemos em que medida podemos falar de agramaticalidade em certas frases quando a sua estrutura não está inteiramente de acordo com as regras da gramática portuguesa. Alguns "angolanismos" são nitidamente marcados e por isso sentidos como incorrectos e pertencentes à linguagem popular. Outros vêm-se tornando comuns, instalando-se podemos dizer que definitivamente na norma, na futura norma do português angolanizado. (*ibid.* 213)

There are two important acknowledgments in Mendes' statement, which had only been implicitly suggested in Marques (1990). First, that the linguistic features that separate AVP from EP, i.e. "angolanismos", are not only lexical but also grammatical. Second, that some of these were so widespread that they could no longer be considered errors but rather part of an emergent standard of Angolan Portuguese[40].

The goal of the works following Marques (1983) and Mendes (1985) in the late 1980s and the 1990s, when the few available studies on AVP were published, was precisely that of identifying the linguistic features that distinguish the emergent Angolan variety of Portuguese from EP at all linguistic levels, be it neologisms and borrowings from Bantu or specific morphosyntactic features.

[40] It can not go unnoticed that both Mendes' acknowledgments are in clear contradiction with the descriptions of AVP depicted in sections 2.1.1, 2.1.2. and 2.1.3., which do not recognize it as a specific standard.

Endruschat (1989), focuses on new word formation mechanisms in AVP and concludes that the variety is typically characterized by three tendencies: preference for shorter and simpler structures, creation of new lexical items based on Portuguese and Bantu words, and transfer from the Bantu languages (p. 73). Despite the fact that Endruschat (1989) draws her data from articles published in the Angolan press between 1980 and 1983, she considers the aforementioned tendencies in AVP to be more frequent in oral rather than written usage; the two are to be treated separately, since "ces deux formes présentent des différences sur tous les niveaux de langue" (*ibid.* 70):

> Les tendances que nous avons décrites sont surtout visibles dans la langue parlée. Dans l'usage official écrit (presse et autres publications), l'apparition de particularités est limitée, pour l'essentiel, au lexique. Les autres niveaux de langue sont caractérisés par une référence consciente à la norme littéraire du Portugal. En d'autres termes, les changements linguistiques sont d'abord vérifies dans la formation de nouveaux lexèmes qui se réfèrent à la réalité angolaise. (Endruschat 1989:73)

In fact, while she rejects the existence of a fully developed creole in Angola, she maintains that, due to Angolans' insufficient proficiency in Portuguese, the Portuguese spoken in the cities shows evidence of being a "secondary creole" (Cf. Valkhoff 1966), e.g. "linguagem dos muceques" (Endruschat 1989:70).

The central role played by oral usage in the development of AVP is reinforced in Endruschat (1990:9-10), unlike the clear-cut distinction between oral and written features of AVP:

> Um grande número de linguistas angolanos (...) e colegas estrangeiros analisaram aspectos da língua portuguesa falada em Angola e todos constataram que se está a desenvolver uma variedade angolana, distinguindo-se por particularidades a todos os níveis do sistema linguístico [...] Todas as particularidades nomeadas ainda não podem ser descritas como formando uma norma angolana do português, mas contudo é possível constatar algumas regularidades na produção oral de textos. Agora o nosso objectivo é demonstrar que estas particularidades entraram parcialmente já na língua escrita.

Moreover, Endruschat (1990) shows that the distinguishing characteristics of written AVP are not limited to the lexicon but include morphosyntactic features too, for the influence of the Bantu languages covers both the lexicon and morphosyntax (*ibid.* 9). Endruschat (1990) puts great emphasis on the role played by the substrate in the development of AVP, and hence she only identifies two of its specific morphosyntactic features (both oral and written), i.e. the use of prepositions and word order of personal pronouns, the last of which is analyzed in greater depth in Endruschat (1993).

A more comprehensive and abundantly illustrated analysis of specific morphosyntactic features of AVP can be found in Gärtner (1989, 1997). The data are also drawn from written texts, but

together with literary examples of the "linguagem dos musseques", examples taken from readers' letters to the editors of several magazines and essays written by school children are also included.

Based on the analysis of his extensive corpora, Gärtner agrees with the hypothesis that AVP's differences from EP originate in Angolans' oral use of Portuguese, especially that of rural Angolans (1997:146), which shows "não raras vezes, traços creolóides" (*ibid.* 146), and that some of the traits of the substandard are now present in colloquial varieties of standard Portuguese (*ibid.* 147). However, unlike Mendes (1985) and Endruschat (1989), Gärtner (1989, 1997) does not perceive these features as part of an Angolan standard of Portuguese in the making, but rather as "desvios do português padrão europeu que têm que ser considerados como fenômenos do substandard" (Gärtner 1997:147).

Another distinctive feature in Gärtner's work is his analysis of AVP's development and synchronic structure not only in comparison to EP but also to BP and Mozambican Portuguese (MP), which leads him to conclude that, despite some minor synchronic details, they have all developed in very similar ways:

> Partimos da pressuposição de que os fenômenos análogos no português do Brasil são devidos a causas análogas. A diferença está em que estes surgiram nos primeiros séculos do Brasil colonial, ou seja, nos séculos 16 a 18, o que fez com que houvesse certas diferenças nos detalhes do resultado lingüístico actual, e que se lhes foi superpondo, a seguir, pelo menos parcialmente e em certas regiões urbanas, um processo de relusitanização. (Gärtner 1997:146)

> Na maioria dos casos é de supor que tenha havido uma *simplificação* do português por falantes de línguas tipologicamente diferentes em situação de *bilingüismo*. Em alguns casos, as alterações daí resultantes coincidem com *tendências inerentes ao português ou às línguas românicas e indoeuropéias*, o que deu margem a diferentes hipóteses sobre a sua gênese. Em outros casos há-de se pensar na *transposição de fenômenos do substandard do português europeu* para o Brasil, Angola e Moçambique. (*ibid.* 172 – *my emphasis*)

Perl (1989b) further elaborates on the hypothesis that Portuguese has developed in similar ways not only in Brazil, Angola and Mozambique, but also in the African countries where Portuguese-based creoles developed. He maintains that the emerging regional varieties of Portuguese in these countries are all a consequence of the social stratification of the language according to the level of instruction of the Portuguese settlers, their social and geographical provenance and the degree of contact they had with African languages. According to him, only a minority of the settlers were educated and proficient in standard EP. The majority of Portuguese speakers, due to different geographical origins and their lack of contact with

Portugal, spoke "une variante de la langue courante" rather than the standard and a large number of uneducated bilingual speakers spoke a "secondary creole" "caractérisé par des interferences plus ou moins fortes, avec la majorité des langues bantoues", e.g. "linguagem dos musseques" (*ibid.* 13).

However, none of these social variants can be interpreted as the standard because "pour l'essentiel la situation du portugais en Afrique est caractérisée par une dépendance des variants diatopiques, des variants diastratiques, provoqués par l'influence permanente de la langue par un nombre élevé de locuteurs bilingues" (*ibid.* 15). Of course, Perl admits regional differentiation according to "les différentes langues africaines de l'adstrat ou du substrat (p. ex. des langues bantoues en Angola et au Mozambique, le creole portugais à São Tomé et au Cap-Vert)" (*ibid.* 14), but maintains that "une variante généralisée d'un portugais africain est concevable étant donné que le développement de la langue est influencé de la même manière dans plusieurs pays car des locuteurs bilingues" (*ibid.* 25).

Hence, Perl (1989b) refers to African Portuguese rather than to specific varieties of Portuguese in Africa, despite the fact that he acknowledges that "la planification du statut du portugais en Afrique est considérée dans tous les pays comme moyen de développement de variantes de langue avec une norme nationale indépendante". In order to define this standard, Perl (1989) warns that "... les facteurs linguistiques des variantes nationales en évolution d'un niveau quelconque du portugais en Afrique ne peuvent être comparés qu'à des éléments corrélatifs de même niveau du portugais européen ou brésilien" (*ibid.* 14)

The hypothesis that there is such a thing as an African Portuguese (AP) can also be found in Vilela (1995, 1999), but unlike Perl (1989b), he seems to use the term to refer to the Portuguese spoken in Angola and Mozambique only:

> O português africano – angolanizado ou moçambicalizado – apresenta indícios claros de afirmação de uma norma própria: (1) na maneira original como adapta o seu vocabulário de origem bantu ao sistema do português – divergindo inclusivamente da norma europeia (lusitana), (2) no modo como simplifica a morfologia flexional do português, (3) como começa a optar pela ordenação dos elementos frásicos na sequência discursiva e, sobretudo, (4) como força o léxico do português a adaptar-se à mentalidade africana (...) o que implica, por vezes, uma reformulação do esquema frásico em alguns dos seus modelos proposicionais. (Vilela 1995:698)

Nonetheless, Vilela (1995, 1999) also agrees that in studying AP "há toda uma gama de situações a destrinçar, de padrões de língua a definir" (1995:49) and identifies four types of AP, which must be treated separately since each represents a different standard: (a) the Portuguese spoken by those in closer contact with the Portuguese or who speak it as a mother tongue (i.e.

very close to standard EP); (b) the Portuguese which sprang from the Portuguese colonization (i.e. colonists from different geographical and social strata); (c) the "secondary creoles" (i.e. *linguagem dos musseques*); and (d) the Portuguese taught in schools and used in adult schooling (1995:49).

Contrary to what one would expect, Vilela (1995, 1999) does not say in what ways these standards differ from one another. In fact, Vilela focuses on what they have in common, pointing out simplification at all linguistic levels (as a consequence of bilingualism and diglossia) as the most distinctive feature of AP and the factor mainly responsible for its growing differentiation from EP, which is similar to the differentiation between Latin and its offspring:

> ... estamos perante a formação de construções típicas, alterações sémicas e classemáticas, enquadramentos e simplificações: os mesmos fenómenos que se verificaram na formação dos idiomas românicos a partir do latim. (Vilela 1995:68)

The major strength of Vilela (1995, 1999) is his abundant illustration of the specific features of AP. The data are drawn from the available literature. In fact, Vilela's work is essentially descriptive in nature, focusing mainly on the synchronic structure of AP:

> Para se poder deduzir com mais segurança as características do português de Angola e Moçambique, formulando princípios sobre interferências, entre essas línguas e a língua portuguesa falada e escrita, há que recorrer a construções frequentes e seus desvios relativamente à norma padrão do português europeu, há que procurar as palavras em domínios que denotam a realidade africana diferente da realidade europeia, etc. (Vilela 1995:52)

Vilela (1995, 1999) shows little interest in investigating the sociolinguistic background responsible for the synchronic data he surveys, which is all the more surprising considering his emphasis on the need to distinguish between different standards of AP according to the social stratification of the language.

Gärtner (1989, 1997), Perl (1989) and Vilela (1995, 1999) have had a lasting impact on the more general literature on non-creole varieties of Portuguese, where it is still possible to come across the assumption that since AVP, BP and MP share important synchronic features, they must have developed in pretty much the same way, which is usually unaccompanied of any analysis of the different sociolinguistic histories of the three countries. However, this line of investigation was not followed by most of the relatively few linguists working on AVP, especially Angolan linguists, who reduced the scope of their works and focused on the description of AVP itself, which they define as a variety in the making. Some of their works consist of the description of specific phenomena of this variety, such as the word order of personal pronouns (e.g. Miguel 1997), gender, tense, mode, aspect and textual cohesion (Costa

1997) or prepositional phrases (Cabral 2005). Others try to offer a more general overview of its synchronic structure, focusing on general trends in phonetics, morphology and morphosyntax (e.g. Mingas 1998, 2000; Barros 2002). In the latter works AVP's specific features are mostly analyzed in connection with the difficulties in teaching Portuguese in Angola.

An interesting aspect of the aforementioned descriptions of AVP is that they distinguish between general features of AVP and specific features according to the region where AVP is spoken. The former include both phonetic, lexical and morphosyntactic phenomena, such as variable rules of number agreement (Mingas 1998; Barros 2002). The latter refer essentially to varying phonetic features depending on speakers' Bantu native language (Mingas 1998; Barros 2002). While these descriptions are too general (none of them is more than 12 pages long), they show that regional variation must be taken into account in a more general description of AVP, which had already been suggested by Mendes (1983).

Unfortunately, the only contribution to our understanding of regional varieties of AVP is Mingas (2000), providing the only book-length description up to now of the lexical, phonetic and morphosyntactic features of a specific variety of AVP, i.e. Luanda Portuguese (LP). According to Mingas (2000), these features are a consequence of the prolonged contact between Portuguese and Kimbundu. Hence, she focuses only on the features that reflect "interferência" and "adaptação" of Kimbundu words and morphosyntactic structures into LP (*ibid.* 23). These features are divided into two groups: lexical features (i.e. loanwords[41]) and morphosyntactic features (i.e. lack of number and gender agreement in the noun phrase, word order of object pronouns, lack of subject agreement, and semantics of the prepositions). The data are essentially drawn from written material (i.e. literary works and some issues of the *Jornal de Angola*)[42].

These features had already been identified as typical of AVP in general, but Mingas was the first to list them for a specific variety of AVP. Moreover, Mingas (2000) not only shows the ways in which the Kimbundo lexicon, sounds and morphosyntactic structures are different from those of Portuguese, but she list some of the processes by which these entered LP. Mingas (2000:78-9) makes the following remark regarding the incorporation of Kimbundu nominal loanwords in LP:

---

[41] Mingas (2000) also includes in this group of features the phonetic adaptation of Kimbundu loanwords to Portuguese.

[42] Mingas (2002:24) states that some of the data are oral sentences she collected during a short stay in Luanda. However, since Mingas does not indicate the source of any of the examples it is impossible to distinguish between written and oral data.

> No que respeita aos nominais, eles podem ser divididos em três grupos. O primeiro engloba os nominais que foram lexicalizados com o respectivo prefixo singular de classe; o segundo reagrupa aqueles que entraram na variante angolana só com a estrutura da sua forma plural e o último integra os nominais lexicalizados sem perda do respectivo prefixo. A variante comporta, por outro lado, nomes que entraram no léxico do português de Angola com os dois prefixos, o do singular e o do plural.

Mingas (2000) is essentially a synchronic analysis of LP, which is also clear in chapters 1 and 3 on the sociolinguistic background of the variety (chapter 3), none of which goes back further than the late 19th century. In fact, as she stated earlier (Mingas 1998:116), this is because she considers AVP to have developed only in the mid-twentieth century:

> ... verificou-se no período colonial[43] uma interferência das línguas maternas africanas no português da época, interferência essa que se agudizou actualmente [...] Com efeito, no após independência, facilitado o contacto mais aberto entre os vários grupos linguísticos, bem como entre o campo e as cidades, devido ao factor guerra, a língua portuguesa atingiu níveis de interferência nunca antes alcançados.

Based on this understanding, Mingas (2002a:45) suggests that there is the strong possibility that a Portuguese-based creole is developing in Angola:

> ... a língua portuguesa, pela sua função veicular, corre o risco de crioulização, devido não só às condições do seu ensino como às situações de contacto linguístico que caracterizam o país [...] a não se verificar uma atitude responsável na solução destes problemas, concorrer-se-á, na melhor das hipóteses, para o aparecimento do fenómeno supra-citado de crioulização.

Barros (2002) is even more explicit, assuming that such a creole already exists:

> A inexistência de estudos científicos sistemáticos (há a realçar somente artigos e trabalhos dispersos) não tem podido dar conta da importância daquela convergência[44] e da riqueza do Português Popular de Angola que se encontra num processo de crioulização e, facilmente, se podem encontrar falantes de pidgin, crioulo, variedades do Português, assim como situações de bilinguismo, diglossia, poliglossia. (*ibid.* 35)

> ... o registo oral/popular angolano tende para a crioulização devido à aquisição deficiente do Português e fossilização dos erros; à ausência de código escrito e à exclusão da vida pública (consequência das desigualdades sociais), criando-se, assim, um novo sistema a partir do cruzamento das línguas em contacto. (*ibid.* 42)

These are strong and innovative statements which, unfortunately, are not backed up by any linguistic data. Nor are the terms *pidgin* and *creole* defined by Mingas and Barros, thus making it difficult to know whether their use of the terms is that of modern contact linguistics or Valkhoff's (1966). Still, Mingas and Barros' statements open new lines for the investigation of

---

[43] It must be noted that the expression "colonial period" refers to the period between the Conference of Berlin in 1885 (when Angola was acknowledged as a Portuguese province) and Angola's independence from Portugal in 1975.

[44] Barros (2002) refers to the convergence between Popular Angolan Portuguese and Portuguese-based creoles.

AVP, which, considering Angola's complex linguistic situation, are not only possible but necessary:

> a) Inventariar todas as situações de contacto linguístico, com indicações referentes não só ao número de línguas em presença, como também à percentagem de locutores das mesmas.
> b) Inventariar o número assim como a complexidade, que caracterizam as situações em que a língua portuguesa está em contacto com outras línguas, permitindo uma informação sobre os diversos tipos de relações existentes entre as várias línguas que coexistem com a portuguesa. (Mingas 2002a:49)

The methodology suggested by Mingas (2002a) implicitly restates the hypothesis that it is the present-day sociolinguistic situation in Angola that will enable us to fully understand AVP's synchronic structure. However, other linguists, such as Lipski (2005), following Valkhoff (1966), argue that this can only be achieved by an analysis of Angola's sociolinguistic background since the early days of settlement to the present. According to Lipski (1995:35), the results of this analysis:

> are consistent with the hypothesis that a stable pidgin and possibly even a creole Portuguese was maintained – perhaps very precariously and in only a few enclaves – in Portuguese southwestern Africa for several centuries, spreading to influence local second-language varieties of Portuguese in major slave-exporting ports such as Luanda and Benguela.

In fact, in referring to the *português dos musseques*, Lipski (1995:16) states that "although musseque Portuguese is a second language for most of its speakers, whose proficiency in Portuguese varies widely, there are a number of recurring traits which lend credence to the notion of a distinctive Angolan form of partially pidginized Portuguese".

However, Lipski has recently dismissed the hypothesis that a creole ever developed in Angola (2005:29):

> Unlike in São Tomé, a creole Portuguese dialect never developed in coastal Angola, principally because the majority of the African population was never displaced from original homelands, and the native languages were never fragmented through forced association with linguistically diverse slave populations. Slaves taken from the interior to be shipped from Angolan ports knew little or no Portuguese prior to arriving in the slaving ports, but acquired some basic skills on the Angolan coast or during the voyage to Brazil.

Of course, this does not dismiss the need for a thorough analysis of the available linguistic and sociolinguistic data on Angola in the past, for many of the features that have been identified as typical of AVP (e.g. lack of number and gender agreement in the noun phrase) also appear in the texts quoted by Lipski (1995, 2005), namely in 15$^{th}$ to 19$^{th}$ century literary imitations of Angolan's Portuguese and late 19$^{th}$ early 20$^{th}$ century reports of Angolans' use of Portuguese.

In fact, the question is not whether one should choose between a synchronic or a diachronic analysis of AVP and its sociolinguistic setting, but rather whether one studies AVP for its intrinsic interest only, or also "as hitherto neglected sources of data on Afro-European linguistic contacts worldwide" (Lipski 1995:35). Unlike most of the studies surveyed here, this thesis will try to analyze AVP from the broader perspective of contact linguistics, the scope, methods and terminology of which will be outlined in the following section (2.2).

## 2.2. Contact Linguistics: an overview

The scientific study of the contact between typologically different languages around the world was only accepted as an independent academic discipline in the mid-twentieth century, and mostly from the perspective of pidgin and creole studies (Holm 1988-89, 2000). Key concepts in the field (e.g. pidgin, creole, restructuring, etc) have been used in various and often controversial ways in the literature, so subsection 2.2.1 seeks to provide their most commonly accepted definitions.

It is only recently that linguists have focused on outcomes of language contact that are neither pidgins nor creoles. Hence, pidgin and creole linguistics has now broadened its scope to become *contact linguistics* (Thomason 1997a; Holm 2004), including the study of a wider range of language contact phenomena, such as those linguistic varieties that share a significant number of structural parallels with pidgins and creoles, but retain enough similarity to their European lexical sources to indicate that they have never been fully creolized. Subsection 2.2.2 offers a summary of the theoretical scope of the discipline as laid down by Weinreich (1953) and more recently by Thomason & Kaufman (1988). Section 2.2.3 discusses the current debates concerning the continuum of language contact outcomes and argues for partial restructuring (Holm 2004) as a coherent theoretical framework to help resolve at least part of them.

### 2.2.1. Key concepts in pidgin and creole linguistics: definitions

The first set of concepts that needs defining, as they are extensively used throughout this thesis and are not exclusive to *contact linguistics*, includes *superstrate*, *adstrate*, *substrate*, *lexifier* and *source/input language* as well as *restructuring*. Hence, a *superstrate language* is typically that of the socially or politically dominant group in a language contact situation. It is often used synonymously with *lexifier language*, i.e. the language that is the source of the majority of the lexicon of the new contact variety. A *substrate language* is that of the socio-politically

subordinate group in a contact situation and an *adstrate language* is one spoken by populations that are neither dominant nor subordinate in the contact situation but nonetheless contribute to the development of the new contact variety. *Source* or *input language* is an umbrella term to refer all the languages that contribute to the development of a new contact variety (Sebba 1997:25-26). *Restructuring* is also an umbrella term to refer to:

> ... all structural modifications that a lexifier language undergoes in the selection and evolution of new linguistic elements, influenced by other, competing languages, in a contact situation. The final outcomes of this process is a new linguistic code which consists of a variable, fairly subtle mixture of both substrate and superstrate features. (Neumann-Holzschuh & Schneider 2000:6)

The term implies that the structure of a language undergoes significant changes, but does not specify what the outcomes of that change are (Sebba 1997:26), as there is no consensus as to what these outcomes or the processes underlying their development may be. For example, *pidgins* are traditionally defined as reduced languages that arise in contact situations involving people who have no shared language and do not learn each other's languages as normally used among native speakers. As they are meant to be used in a limited range of contexts, e.g. trade, work or interethnic communication, they are usually no one's mother tongue; the lexicon is extremely reduced and mainly drawn from a single source, and their grammar derives from the simplification of the different input languages, e.g. no definite word order, reduced inflectional and derivational morphology, reduced number of grammatical words and simpler phonological systems than their source languages (Holm 1988:4-5; Arends, Muysken & Smith 1995:17-39; Sebba 1997:37-69).

However, some languages traditionally defined as pidgins (e.g. Tok Pisin) have acquired native speakers (i.e. are nativized) and consequently have stabilized and expanded both their social functions and linguistic resources[45]. These apparent exceptions have traditionally been dealt with in two ways. Some, like Thomason (2001:161), assume that there are prototypical pidgins, that conform to the definition above of languages "restricted in functions and in linguistic material", and deviations from that prototype. Others, like Mühlhäusler (1997), assume that these languages are the intermediary stage of a pidgin developmental continuum that ranges from a pre-pidgin, with no fixed rules and no native speakers (i.e. jargon stage) to an expanded pidgin to a creole.

[45] There is no consensus on the nature of the process(es) leading to the stabilization of pidgins, but Whinnom (1971) suggests that it requires speakers of different substrate languages to adopt the pidgin for communicating with each other, i.e. *tertiary hybridization*.

The idea that the nativization of a pidgin and consequently its social and linguistic stabilization and expansion are key factors in the development of creoles has been opposed by several linguists, who claim that social and linguistic expansion can occur without or before pidgins undergo nativization, that nativization does not always imply significant linguistic change, e.g. Tok Pisin (Thomason 2001:161; Samarin 1997) and that not all creoles evolve from a pidgin, e.g. abrupt creoles (Thomason & Kaufman 1988).

This controversy derives ultimately from the fact that different linguists have different theories concerning the genesis of pidgin and creole languages. These date back to the late 19th century and have been analyzed at length in several introductions to the field (Holm 1988, 2000; Arends, Muysken & Smith 1995; Sebba 1997 and Mühlhäusler 1997). Hence, it will suffice to say here that the most widely debated theories are the *baby-talk hypothesis* (i.e. creoles reflect the simplified version of the lexifier languages used by Europeans when talking to their slaves and servants), the *superstratist hypothesis* (i.e. creoles are essentially nonstandard European dialects only slightly influenced by the substrate – e.g. Chaudenson 1979), the *monogenetic hypothesis* (i.e. all creoles derive, either directly or by means of word for word translation, from the 15th century Portuguese-based pidgin used in trade as a lingua franca along the coast of Africa), the *substrate hypothesis* (i.e. creoles are a combination of European lexicon and the grammatical structure of the substrate languages – e.g. Adam 1883, Taylor 1971,1977; Alleyne 1980), and the *universalist hypotheses* (i.e. the specific features of pidgin and creole languages cannot be explained exclusively by reference to similarities among the input languages but rather by linguistic universals[46]).

This disagreement results also in different theories concerning issues such as the role played by children versus adults in the formation of pidgins and creoles, the length of time needed for these languages to develop, and the features, both linguistic and social, if any, that distinguish them from other languages. For example, Bickerton (1984:173 quoted in Winford 2002:289-90) argued that "the innovative aspects of creole grammar are innovations on the part of the first generation of children who have a pidgin as their linguistic input, rather than features transmitted from pre-existing languages", but this view has been discredited by recent sociohistorical evidence that shows that adults, and not children, were the key agents in creolization. Concerning the length of time needed for creoles to develop, the positions range from those who claim that creoles can arise abruptly (e.g. Bickerton 1981, 1984; Thomason &

[46] The supporters of the later hypotheses diverge on whether the linguistic universals that play a role in the development of creoles are first (Bickerton 1984) or second-language acquisition universals (Thomason & Kaufman 1988).

Kaufman 1988) and those who support more gradualist theories, according to which creoles arise during an extended period of time (e.g. Chaudenson 1992; Arends 1993). Concerning the differences between creole and non-creole languages, Bickerton (1981) and McWhorter (1998) have suggested that the former differ from the latter in specific sets of linguistic traits. However, the generally accepted view is that, while it is possible to find features that seem to be more common in creole rather than non-creole languages (e.g. the use of free rather than bound morphemes), the differences between the two is not their linguistic structure but rather their sociolinguistic histories. The same applies to differences between creoles (Winford 2003:308):

> From this perspective, creoles are simply contact languages that emerged primarily in plantation settings in various European colonies throughout the world. Such settings shared a number of sociopolitical and demographic characteristics, including the use of large numbers of slaves who were transplanted from their homelands and placed under the control of a small minority of Europeans. [...] ... differences in the social settings of each colony led to diversification in the outcomes of the contact between Europeans and the oppressed groups.

In recent years, linguists have achieved great progress in the study of the sociolinguistic settings of the different creoles around the world. The sociolinguistic factors that have been found to be relevant in their development include demographics, degree of mortality in the plantation colonies, nature of community settings, codes of social interaction, duration of contact and degree of bilingualism (Winford 2003:310-14).

It is clear from the discussion above that no single theory can account for the origin and synchronic structure of creoles as these languages display immense diversity. In fact, the majority of linguists today support the idea that the answer lies in the combination of all these theories. Hence, a comprehensive definition of creole could be the following:

> A mixed language that is the native language of a speech community. Like pidgins, creoles develop in contact situations that typically involve more than two languages; also like pidgins, they typically draw their lexicon, but not their grammar, primarily from a single language, the lexifier language. Some creoles arise as nativized pidgins, some arise abruptly with no pidgin stage, and others arise gradually, with or without pidgin stage. Crucially, the creators of a creole (unless it is a nativized pidgin) are not bilingual in their interlocutors' languages. (Thomason 2001:262)

### 2.2.2. Contact linguistics: the theoretical foundations

The broadening of pidgin and creole linguistics to *contact linguistics* results from the general agreement today that the origin and synchronic structure of pidgins and creoles can only be fully understood from the perspective of a wider theory of language contact. A simple, straightforward definition of the discipline is that it is the sub-field of linguistics that seeks "to

study the varied situations of contact between languages, the phenomena that result, and the interaction of linguistic and external ecological factors in shaping these outcomes" (Winford 2003:5), at the level of both the group and the individual who is bi- or multilingual.

Contact linguistics has only recently been established as an independent academic discipline, but the scientific study of language contact dates back to the late nineteenth century (*ibid.* 8), when there was a surge in the study of contact vernaculars as a means to counter-evidence the traditional claims in comparative historical linguistics that language change was internally-induced only and that hence language mixture was impossible. Schuchardt, in a series of papers written between 1880 and 1914, was particularly important in this regard (Winford 2002:288). Research focused on a wide range of topics (e.g. types and mechanisms of cross linguistic interference; creation of new contact languages such as pidgins and creoles; constraints on language mixture, etc), using both linguistic and social criteria. However, the first comprehensive framework to study language contact was not provided until Weinreich (1953).

Weinreich (1953) emphasized the need for a "broad approach" (*ibid.*111) in contact linguistics, which he termed "interference studies", that combined the analysis of not only the mechanisms and structural causes of language contact at all linguistic levels, but also of psychological theories of bilingualism and the socio-cultural setting of language contact because

> ... certain forms of interference are called forth, facilitated, or inhibited by concrete structural differences of the languages. But (...) the total impact of the languages on each other – the extent to which interference of each structurally determined type is manifested – can hardly be accounted for by strictly linguistic data. [...] For practically every form of interference, there is an interplay of factors external to the structures of the languages which favor or inhibit the development of interference of that type.

Hence "to predict typical forms of interference from the sociolinguistic description of a bilingual community and a structural description of its languages is the ultimate goal of interference studies" (*ibid.* 86). Weinreich (1953) emphasized the unity of all contact phenomena but predicted two types of contact possibilities: *interference* and *language shift*. The former implies "the rearrangement of patterns that result from the introduction of foreign elements into the more highly structured domains of language, such as the bulk of the phonemic system, a large part of the morphology and syntax, and some areas of the vocabulary" (*ibid.* 1)[47]. This rearrangement may crystallize and lead to the development of new languages such as pidgins and creoles. The criteria that contribute to the development of these new languages are both linguistic (i.e. degree of difference between the languages in contact) and social (i.e.

[47] Cf. the definition of *restructuring* given at the beginning of section 2.2.1.

stability of form, diversity of functions and speakers own perceptions). Hence, pidgins and creoles are defined by Weinreich (1953:69) as languages that:

> [...] have attained some or all of the following: (1) a form palpably different from either stock language; (2) a certain stability of form after initial fluctuating; (3) functions other than those of a workaday vernacular (e.g. use in the family, in formalized communication, etc.); (4) a rating among the speakers themselves as a separate language.

On the other hand, language shift, i.e. "the change from the habitual use of one language to that of another" (*ibid.* 106), is "entirely extra-structural", since "the respective structures of the two languages in contact never determine which language is to yield its functions to the other" (*ibid.* 107). Even so, linguistic factors are still important in studying language shift for depending on a series of social factors (e.g. domains and functions of language use, social status, language loyalty, duration of contact, etc) "the speakers of the losing language may learn a new language so well as to leave no trace on it" or "they learn it in an imperfect manner, bequeathing the phonetic and grammatical peculiarities of their speech to future generations in the form of a substratum" (*ibid.* 109).

However, Weinreich (1953) offers "no means of predicting what types of contact-induced changes will occur when" (Thomason & Kaufman 1988:13). According to the latter authors, this is because he believed that ultimately "the structure of a language determines what can happen to it as a result of outside influence" (Thomason & Kaufman 1988:14). Thomason & Kaufman (1988) agree that linguistic constraints are important in language contact, but they warn that they have "no value for the development of a predictive theory of linguistic interference" (*ibid.* 34) because "it is the social context, not the structure of the languages involved, that determines the direction and the degree of interference" (*ibid.* 19).

In fact, Thomason & Kaufman (1988) were the first to provide an overall predictive theory of contact-induced language change, which is today the most widely accepted in the field. Their goal was to "describe and analyze linguistic results of language contact situations, and to correlate these results with certain fairly general kinds of social factors" (*ibid.* 36). Hence, they draw a distinction between two different types of language contact situations, i.e. language maintenance and language shift, which are associated with two different mechanisms of interference: borrowing and interference, respectively.

Thomason & Kaufman (1988:37) define *borrowing* as "the *incorporation* of foreign features into a group's native language by speakers of that language: the native language is retained but is changed by the addition of the incorporated features". *Interference via shift* is the result of

"imperfect group learning during a process of language acquisition" (*ibid.* 38) in which "the errors made by members of the shifting group in speaking the target language then spread to the target language as a whole when they are imitated by original speakers of that language" (*ibid.* 39). Interference via borrowing starts with the lexicon (though structural features may be borrowed under strong long-term cultural pressure from source-language speakers), whereas interference via shift begins with phonology and syntax, sometimes including morphosyntax and it is initiated by the shifting speakers, not the original target language (TL) speakers. However, Thomason & Kaufman (*ibid.* 45) warn that the two types of interference are not mutually exclusive as "target language speakers may be borrowing words and possibly even structural features from a language whose speakers are in the process of shifting to the target language and incorporating their learner's errors into it".

The relevant social factors in both language maintenance and shift are the relative population sizes and the intensity of contact, i.e. length of contact and the level of bilingualism, which Thomason & Kaufman correlate with different degrees of interference. For language maintenance Thomason & Kaufman (1988:74-6) define a five-level borrowing scale: (1) casual contact: lexical borrowing only; (2) slightly more intense contact: slight structural borrowing (e.g. conjunctions and adverbial particles); (3) more intense contact: slightly more structural borrowing (e.g. adpositions, derivational affixes; inflectional affixes in borrowings; personal and demonstrative pronouns); (4) strong cultural pressure: moderate structural borrowing (e.g. extensive word order changes, borrowed inflectional affixes added to native words); and (5) very strong cultural pressure: heavy structural borrowing (e.g. changes in word structure rules, addition of concord rules).

Concerning language shift, Thomason & Kaufman (1988) distinguish between language shift with normal transmission and language shift without normal transmission. The former comprehends three degrees of interference: (1) shift without interference (i.e. small shifting group relative to the numbers of TL speakers, slow shift and consequent bilingualism), (2) slight interference (i.e. essentially phonological and syntactic interference), (3) moderate to heavy interference, not only in phonology and syntax, but also in morphology (i.e. large shifting group, little access to the TL and consequent imperfect learning)[48]. In language shift without normal transmission shift occurs rapidly and the shifting group is so large numerically that the

[48] The interference phenomena identified by Thomason & Kaufman (1988) in shift situations affect the phonology (e.g. phonemic mergers or replacement of more marked by less marked features; elimination of allophonic alternations), syntax (e.g. changes in the word order) and morphology (e.g. mergers of categories or their replacement by syntactic expressions – e.g the loss of gender).

TL model is not fully available to all its members, which leads to successful acquisition of the vocabulary only. The possible outcomes in shift without normal transmission are abrupt creoles, pidgins and bilingual mixed languages.

Departing from the standpoint of historical linguistics, Thomason & Kaufman (1988) argue that the difference between *shift with normal transmission* and *shift with abnormal transmission* is that in the former, in spite of the structural interference features "enough inherited grammatical patterns remain that genetic continuity has clearly not been disrupted" (p. 129), whereas in the latter the grammatical structure of the resulting languages cannot be traced back to a single parent language.

The reason why Thomason & Kaufman (1988) are so important in contact linguistics is because they provided the theoretical framework that enabled linguists to realize that "the processes of change and restructuring that give rise to pidgins and creoles are different only in degree rather than kind from those that occur in many cases of language contact" and hence, rather than attempting to define "pidgins and creoles in relation to other languages", linguist should focus on attempting to define "pidgins and creoles in relation to other kinds of contact languages, placing them within a typology of contact outcomes" (Winford 1997:2-3).

However, the typology of contact outcomes proposed by Thomason in later research (1997, 1997a) is not unproblematic, for it only comprehends three classes, i.e. pidgins, creoles and bilingual mixed languages:

> a *contact language* is a language that arose by some historical process other than normal transmission. Or, to put it in another way, a contact language is comprised of grammatical and lexical systems that cannot all be traced back to a single parent language. (Thomason 1997:75)

According to Thomason (1997a:3):

> there are prototypical contact languages in all three categories, and there are also various kinds and degrees of deviations from the prototype. All these deviations (e.g. "semi-creoles" and "koinés") are best analyzed in relation to the prototype, as on a continuum, and not as separate types of contact languages.

Thomason (1997:86) acknowledges that one must situate these deviations "with respect to the three main classes by showing how they differ from prototypical contact languages", but she states that "there are no congruent social and linguistic correlates by which they may be identified" (*ibid*). What Thomason (1997, 1997a) seems to overlook is that while borderline cases such as semi-creoles may indeed not be classes of contact languages on their own, some orderly generalizations can be made regarding their development and synchronic structure, as

the following section will hopefully show. Moreover, such a blunt statement about the impossibility of establishing "congruent social and linguistic correlates" by which these borderline cases may be identified strikes us as surprising at a time when case studies on these varieties are still in their infancy.

### 2.2.3. Continuum of contact languages and partial restructuring

It was only in the 1960s that linguists first suggested a theory to account for the origin and synchronic structure of those contact varieties that display a significant portion of the European languages' morphosyntax as well as the introduction of substrate and interlanguage features, i.e. one of the borderline cases identified by Thomason (1997, 1997a). As summarized by Holm (2000:10), it was assumed that due to continued contact with the lexifier language, the creole would gradually decreolize, i.e. lose its creole features, leading to a continuum of creole varieties:

> from those farthest from the superstrate (*basilect*) to those closest (*acrolect*), with *mesolectal* or intermediate varieties between them. After a number of generations some varieties lose all but a few vestiges of their *creole features* (those not found in the superstrate) through decreolization, resulting in *post-creole* varieties such as (according to some) African American Vernacular English or Brazilian Vernacular Portuguese.

By the end of the 1970s, *decreolization* alone was generally agreed to explain why creoles displayed varying degrees of restructuring (Holm 2004:7). This view came to be questioned in the 1980s as there was growing evidence that (1) not all creoles that remained in contact with the lexical source language had undergone decreolization (cf. the case of Palenque, Colombia), and (2) not all features traditionally defined as acrolectal in the creole continuum model were necessarily the product of gradual decreolization but might rather have existed in the variety since the early days of its development. Hence, referring to contact varieties such as African American English, Afrikaans and Brazilian Vernacular Portuguese as *post-creoles* was a misuse of the term, as they may have never undergone full creolization.

Over the last two decades, supporters of different theories of pidgin and creole genesis have tried to provide explanations to account for the development and synchronic structure of these varieties, but the most comprehensive theoretical model suggested so far is *partial restructuring*[49], which was developed and re-formulated by Holm in several papers since the

[49] Building on the work of early creolists such as Coelho (1880:43-4; 1886:160,170) and Schuchardt (1888:67), the theory was first discussed by Holm under the name *semi-creolization*. For a detailed account of the works

late 1980s, culminating in a recent book-length comparative analysis of five of such partially restructured vernaculars (PRV), i.e. Afrikaans, African American English, Brazilian Vernacular Portuguese, non-standard Caribbean Spanish and Réunionnais (Holm 2004).

The major strength of Holm (2004) is that, in comparing five different PRVs, he is able to provide an overview of the general structural and sociolinguistic tendencies of the languages that have undergone partial restructuring. The theory shows how particular sociolinguistic factors (i.e. demographic balance of native versus non-native speakers of the European source language during the first one hundreds years of contact and the length of contact) and particular linguistic processes (i.e. language drift, leveling, imperfect language shift and language borrowing) combine to produce the specific structure of these PRVs (i.e. a contact variety characterized by the retention of a significant portion of the European languages' morphosyntax as well as the introduction of substrate and interlanguage features).

Hence, according to Holm (2004), what distinguishes PRV from their lexical source languages and the creoles and post-creoles of the same lexical base is the fact that PRVs emerge in a sociolinguistic setting in which, at the early period of their development, there is a more balanced ratio of native vs. non-native speakers of the lexical source language in a community whose populations speak different mother tongues[50] and are in the process of changing to a typologically different target language to which learners have only partial access due to social restrictions often connected to servitude. Under these sociolinguistic conditions the following linguistic processes can occur. First, the European language can undergo *language drift* and as its different regional and social varieties come into contact it can undergo *primary leveling*. Substrate speakers then imperfectly shift to the TL. *Imperfect language shift* is the key process in partial restructuring as shifting speakers pass on structural features of ancestral languages and interlanguages to their monolingual descendants. This is followed by *borrowing* of substrate features, but due to continued contact with the lexifier source language *secondary leveling* occurs, leading to the loss of many of the features not found in the TL and consequently to a greater proximity with the latter.

Partial restructuring theory is not so much a claim in favor of a new type of contact language (cf. Thomason 1997), but an attempt to make an orderly generalization about the origin and synchronic structure of these segments in the continuum of language contact that have fallen

---

of these creolists see Holm (2004:1-23). For a first-hand explanation of the reasons that led Holm to replace *semi-creolization* with *partial restructuring* see Holm (2004:xi-xviii).

[50] The fact that partial restructuring occurs among a population speaking a number of different first languages distinguishes it from language shift by linguistically homogeneous groups, bringing it closer to full creolization (Holm 2004:142).

between "the cracks of theory, being neither unrestructured overseas varieties nor fully restructured creoles" (Holm 2004:144). And the reason why it is so helpful in increasing our understanding of the range of possible outcomes of language contact is because it shows that "there is no need to choose one linguistic process over the other in accounting for the structure of the new varieties" (*ibid.* 141). In short, it accepts the existence of fuzzy boundaries in language contact and provides a coherent theoretical framework to account for at least some of them.

## 2.3. Summary and Conclusions

The brief overview of the literature on AVP (Section 2.1) shows that it is one of the least studied non-European varieties of Portuguese. In fact, research on AVP is extremely recent. Early Portuguese grammarians refer to language contact phenomena in Angola as early as the 16th century and Schuchardt (1888) provides evidence of the development of a distinctive restructured variety of Portuguese in the country by the 19th century. However, early Portuguese dialectologists and contemporary grammars of Portuguese make no reference whatsoever to the existence of a distinctive Angolan variety of Portuguese. The emphasis has been put on the worldwide structural uniformity of Portuguese, allowing only for the description of minor phonetic and lexical contact-induced specificities (e.g. Martins 1958).

Only following Angola's independence from Portugal in 1975 was there growing awareness among linguists, especially those in Angola, that a new contact vernacular in the making was spreading throughout the country as either a first or second language of the population, i.e. AVP. The few available works on this new variety consist mostly of enumerations of the features that distinguish it from EP, whether neologisms and borrowings from Bantu languages (Endruschat 1989; Mingas 2000) or specific morphosyntactic features (Endruschat 1990; Gärtner 1989, 1997; Vilela 1995, 1999). The data used in these works is not unproblematic, in that it is mostly drawn from written corpora (e.g. literary works, Angolan newspapers and magazines, and letters).

Moreover, none of the works surveyed in section 2.1 distinguish between features developing in the early days of contact and contemporary features, or between the AVP spoken in urban coastal cities (where the presence of L1 Portuguese is stronger) and that of the interior (where Bantu languages are dominant). Sociolinguistic factors relevant in language contact situations are neglected although their relevance is referred to (e.g. the extent of displacement of the

speakers of the relevant substrates; the degree of learners' access to the TL; the nature and extent of contact; the ratio of native versus non-native speakers of Portuguese).

Nonetheless, the overview of the literature on AVP in section 2.1. suggests that, apart from the suggestion that AVP does not differ from EP except in its oral use (essentially at the phonetic and lexical levels, i.e. neologisms and borrowings from Bantu - cf. Boléo 1955; Carvalho 1971; Cuesta & Luz 1971; Teyssier 1997; Mira Mateus *et al* 2003), there are at least two hypothesis to consider in accounting for the development and synchronic structure of general AVP. One hypothesis is that the synchronic features of AVP are simply the remnants of a 15th century Portuguese-based pidgin used as lingua franca in coastal Africa until the 19th century, which due to growing contact with Portuguese developed into a variety closer to the standard but with sporadic deviations from it (Cf. Valkhoff 1966; Lipski 1995, 2005). An alternative hypothesis, suggested by the majority of the works surveyed above, is that AVP never underwent full restructuring. Its creole-like features derive from the natural drift of Portuguese, dialect leveling and transfer from Angolans' L1 due to multilingualism and imperfect language acquisition, especially after 1975, when a mainly Bantu-speaking population had to acquire some degree of proficiency in Portuguese. Hence, AVP's synchronic structure may be essentially a 20th century phenomenon.

The overview of the literature on contact linguistics in section 2.2, especially in subsection 2.2.3 shows that these two hypotheses are not mutually exclusive and can be accommodated under the theory of partial restructuring. AVP's similarity to BVP, i.e. its retention of a significant portion of the morphosyntax of EP, as well as the introduction of substrate and interlanguage features, seems to indicate that this hypothesis is worth careful investigation, despite the fact that while in Brazil the necessary conditions for partial restructuring to occur were present during the early days of settlement, in Angola this process is still going on, especially in the interior.

# Chapter 3

## AVP: the sociolinguistic setting of its development

### 3.0. Introduction

As shown in the previous chapter, it is widely accepted in contemporary contact linguistics that "every outcome of language contact has been associated with a particular kind of social setting and circumstances that shape its unique character" (Winford 2003:10). Consequently, any attempt to understand the synchronic structure of a given contact variety must not only "describe and analyze the linguistic results of language contact" but also "correlate these results with certain fairly general kinds of social factors" (Thomason & Kaufman 1988:36). The demographics of the populations in contact and the intensity of that contact (i.e. the length of contact and the level of bilingualism), as well as the social patterns of colonization and how these affected the nature and extent of language contact, the displacement of the substrate speakers, and the degree of learners' access to the target language have been identified as key social factors in shaping the outcome of language contact, despite the fact that linguists are not yet agreed on the best ways to quantify the relative weight of each of these factors (Winford 2003:25).

Understanding how specific social factors combine to produce a given linguistic structure is particularly important in the case of those language varieties such as AVP, which share a significant number of structural parallels with pidgin and creole languages, but retain enough similarity to their European lexical sources to indicate that they have never been fully creolized. As shown in Holm (2004), it is the specific sociolinguistic setting at the early period of their development and the linguistic processes it triggers that distinguish partially restructured vernaculars from their lexical source languages and the creoles and post-creoles of the same lexical base.

Building on this theoretical approach, the purpose of this chapter is to analyze the existing social setting in Angola during the six relevant periods of the history of Portuguese presence in the country (i.e. sections 3.1 to 3.6) in order to evaluate how particular sociolinguistic factors combined to produce the specific structure of contemporary AVP (i.e. section 3.7).

## 3.1. Portuguese overseas expansion until the discovery of the Congo

The Portuguese overseas expansion can be said to have been motivated by "(i) the crusading zeal against the Muslims, (ii) the desire for Guinea Gold, (iii) the quest for Prester John, (iv) the search for Oriental spices" (Boxer 1997:18). By 1482, the Portuguese had achieved the two first goals, securing key positions in the North of Africa against the Moors and establishing *feitorias* along the coast of West Africa as far as the Gold Coast which were used as trading posts for gold and slaves. Gold was used to sponsor further journeys in search of India and Prester John[1], whereas slaves were taken to Europe or used in the settlement of the previously uninhabited Atlantic islands (e.g. Madeira, Açores, Cape Verde, São Tomé, etc).

However, once the transport and the profitability of West African gold and slaves was secured, the main goal of the Portuguese during the first stage of their overseas expansion (c. 1419 to 1499) was to find the maritime route to India rather than to increase their presence on the African continent (Boxer 1997:15; Riley 1998-2000:160-1). Consequently, only after Vasco da Gama returned from India in 1499 did the Portuguese crown focus on consolidating its possessions on the coast of West Africa, but only with the purpose of securing the acquisition of more slaves to take to India and later to Brazil. Hence, until 1499 the Portuguese presence on mainland Africa was limited to the few forts on the coast which supplied a growing slave trade (Riley 1998-2000:161).

The sociolinguistic settings that resulted from this overseas expansion policy were different in the different areas of the African mainland. In Upper and Lower Guinea, where Portuguese-based creoles developed, as well as on the Gold Coast, slaves, gold and ivory were the principal commodities sought by the Portuguese, most of whom were exiled criminals (also called *lançados*) who often married African women. These Portuguese and their mixed-race children (*tango-maos*) functioned as principals or intermediaries in the trade, often penetrating a considerable distance into the interior and establishing Portuguese as the lingua franca of these regions (Boxer 1963:9-11). In Upper and Lower Guinea and on the Gold Coast "the Portuguese relied not only on peaceful contacts but on a display of power and force, as exemplified by the erection of the castles at São Jorge da Mina (1482) and Axim (1503)" (*ibid.* 11).

The situation in the Congo, where Diogo Cão first arrived in 1482, was significantly different, especially concerning the social background of the Portuguese sent there and the type of relationship that was established between the Portuguese crown and its Congolese counterpart.

[1] For a survey of the reconnaissance journeys carried out by the Portuguese in mainland Africa see Santos (1988).

This took on a diplomatic and institutional character that was to be unique in the history of Portuguese West Africa. The sociolinguistic consequences of this new expansion policy are analyzed in section 3.2 below.

## 3.2. The early contacts (1482-1565)

### 3.2.1 Congo Kingdom

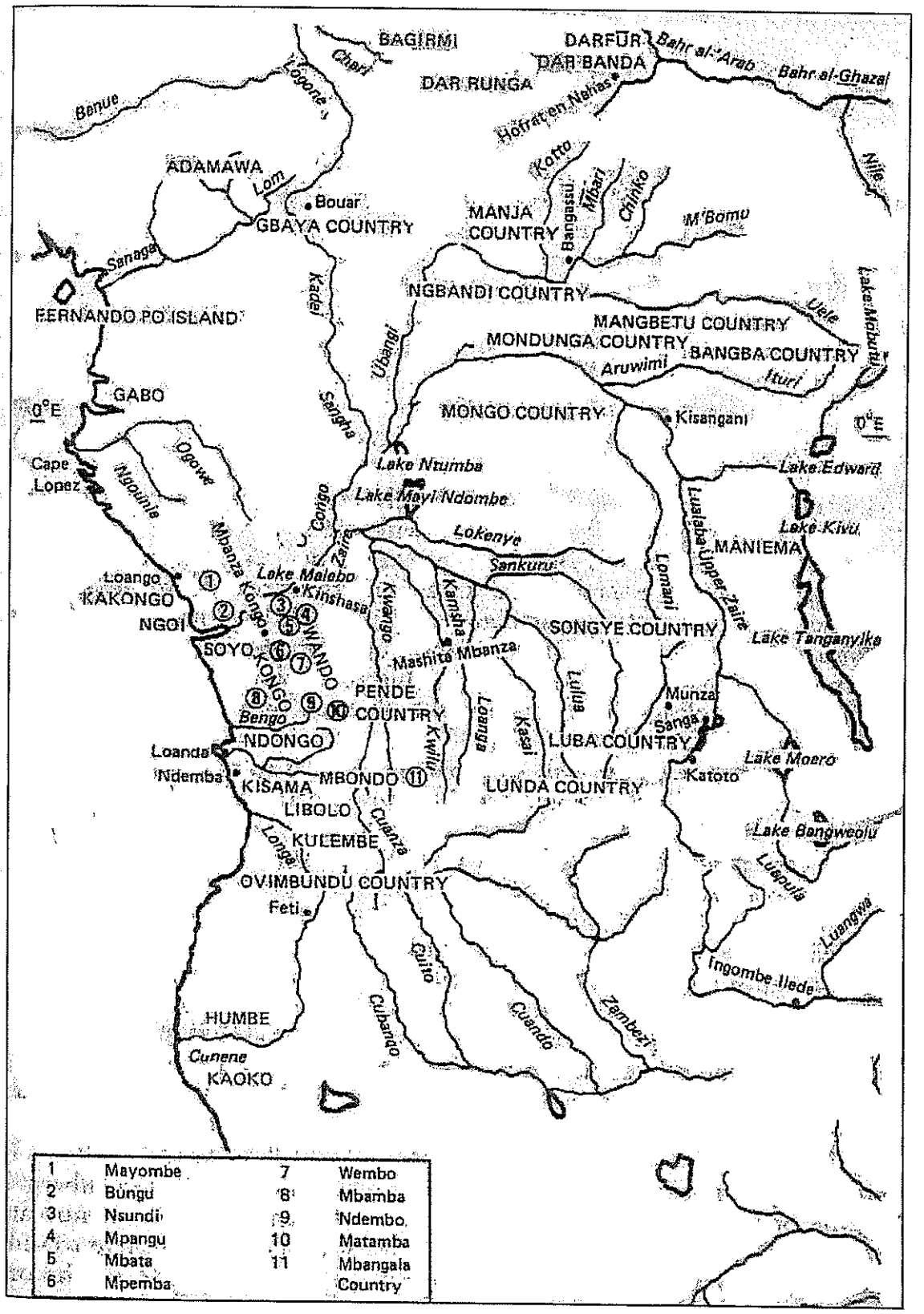
On arriving at the regions south of the Gold Coast in search of the sea route to India in 1482, the Portuguese found a constellation of powerful kingdoms, the most powerful of which was the Congo kingdom, which was ruled by *Nzinga a Nkuwu*[2]. It was densely populated; it had a highly structured society consisting of nobility, villagers and slaves, a very centralized government with a military and its own currency, called the *nzimbu* (Vansina & Obenga 1992:274-276; Iliffe 1999:109). The extensive sphere of influence of the king included not only the six provinces of the kingdom (i.e. Mpemba, Nsumdi, Mpangu, Mbata, Mbamba and Soyo) but also other Bantu-speaking kingdoms, such as the Ndongo, which paid tribute to him (Birmingham 1966:4). The capital of Congo was *Mbanza Congo,* located in the interior province of Mpemba. Its population "was estimated to be as much as 100,000 people" (*ibid.* 2)[3]. Kikongo was the language of the bulk of the population but it is likely that other related Bantu languages were also spoken, especially among the slaves taken from neighboring enemy kingdoms (cf. Map 4ed).

The Portuguese crown quickly realized that an alliance with such a powerful kingdom could be of great help in their search for Prester John and in their further exploration of mainland Central Africa (Santos 1988:45-53). In other words, diplomacy and institutional cooperation had more to offer than war and conquest. Hence, a bilateral agreement was established between the two crowns during the reign of John II of Portugal (c. 1481-1495) and John I of the Congo (c. ?-1509 ), which was later reinforced during the reigns of Manuel I of Portugal (c. 1495-1521) and Afonso I of the Congo (c. 1509-1543). The Congo was to receive Portuguese missionaries, political counsellors and specialized workers (e.g. masons, bakers, teachers, doctors, etc) and Portugal was to have the opportunity to spread the Catholic faith and acquire African slaves and goods such as gold and ivory (Russel-Wood 1998-2000a:242-3).

[2] Other powerful kingdoms in the region were Loango, Lunda, Luba and Matamba. For a brief history of these kingdoms or states see Birmingham (1966), especially pages 1 to 20.

[3] For a detailed description of *Mbanza Congo* see Amaral (1987a).

**Map 4: Central Africa, c. 1500**
**(Vansina 1997:220)**



The diplomatic and institutional nature of this agreement had an important effect on the social background of the Portuguese that were sent to the Congo and the type of activities they were involved in as well as their sphere of influence in the kingdom, and it also affected power relations between them and the native rulers. This contributed to a sociolinguistic setting that was different from that in Upper and Lower Guinea, where Portuguese-based creoles developed.

In fact, the few Portuguese that were sent to Congo were expected to act not as traders but as diplomats, providing political, military and religious assistance to the king. Consequently, due to an order of Manuel I in 1512, the Portuguese sent to the Congo did not include criminals or convicts, but only missionaries and noblemen, some of whom married women of the Congolese nobility with their children entering the line of succession (Russel-Wood 1998-2000a:242). The majority of the Portuguese were accommodated at the king's court in Mbanza Congo in European-style houses especially built for them (Amaral 1987a:8). The only area where the Portuguese were allowed to settle other than the capital was at the port of Mpinda, where they built a fort to receive the slaves that were captured in the interior by the Congolese and who were to be sent to Portugal and São Tomé. In fact, the control of the territory's justice system and trade remained fully under the control of the king of the Congo, who controlled the movements of the Portuguese according to his own interests. For example, none of the expeditions that the Portuguese tried to make in the Congo in order to explore the Congo River and find Prester John received the necessary support from Afonso I (Santos 1988:49). Power relations between Portugal and Congo were not as unbalanced as in other regions in Africa, where the Portuguese simply imposed their will on the native chiefs.

Moreover, with the exception of common slaves, the Africans taken to Portugal were mostly members of the Congolese nobility. In Lisbon, they were treated according to their social and political status, receiving expensive gifts and education, learning the basic principles of the Catholic faith and participating as guests of honour in the Portuguese court (Curto 1998-2000:414). It is also interesting to notice that these noblemen were often young boys, whose proficiency in Portuguese was bound to be much better than that of the four adult noblemen who were taken to Lisbon by Diogo Cão earlier in 1482. Hence, unlike the Upper and Lower Guineans, the Central Africans taken to Europe included youngsters sent to learn the basics of Portuguese so as to be useful as interpreters in trade. They also included many young nobles who were to serve diplomatic functions between the two allied states and who consequently had to acquire good mastery of both spoken and written Portuguese and Latin.

Hence, Portuguese influence in the Congo was significant, but owing to their reduced numbers, their social and political background, their establishment in Mbanza Congo rather than in other

parts of the kingdom and the fact that they had no jurisdiction over the kingdom's politics, commerce and army, their influence was restricted to what they could exert on the ruling classes (i.e. courtiers, provincial governors, etc), especially concerning clothing and titles of nobility (Santos 1988; Amaral 1987a). Of course, many Portuguese loan words entered Kikongo (cf. Martins 1958; Bal 1975, 1979), but it is unlikely that Portuguese was ever widely used throughout the kingdom in contexts other than official correspondence between the Portuguese and Congolese kings and between the latter and their provincial governors, many of whom were the mixed-race children of the Portuguese.

In fact, there were very few reasons for any widespread use of Portuguese in the Congo. The Portuguese missionaries themselves either had very little contact with the general population as their main function was to help create a local clergy, or if they did, they adopted African languages. Among the ruling classes, communication was most often established by means of interpreters. This does not mean that the Congolese nobility had no knowledge of Portuguese, but it is important to notice that what they were exposed to was either European Portuguese (i.e. as spoken by the Portuguese among themselves) or a variety very similar to it, as suggested by the accuracy of the language used by the European-educated interpreters and secretaries in the aforementioned correspondence. In fact, according to Oliveira (1536:45), the Africa-born speakers of Portuguese perceived their speech as similar to that of their European counterparts, which is consistent with the hypothesis that the two varieties were quite similar. Moreover, the linguistic traits that were identified by Oliveira (1536) and Barros (1540) include phonological phenomena only, all of which could also be found in 16th century EP (cf. chapter 2, section 2.1.3 for discussion of these features). Finally, slave traders were the least likely to need Portuguese as they operated essentially in the Kikongo-speaking provinces of the kingdom or in neighboring countries where closely related Bantu languages were spoken.

The language used by the lower social classes is not documented, but the socio-historical setting in the Congo suggests that they were likely to have continued to use their own languages as they had very little contact with the Portuguese. Moreover, that contact was usually of a religious nature and was with the local Kikongo-speaking clergy. The fact that the first catechisms in Bantu were published during this period (cf. chaper 2, section 2.1.4) also indicates that it was the African languages and not Portuguese that was widely used in Christianization, even when the missionaries involved were Portuguese.

Lipski (2005:63) suggests that the common people "possessing only a passing acquaintance with Portuguese would speak a rough pidgin, similar to that found in rural regions of contemporary Angola. Their language does not appear in documents of the time, but it is

unlikely that this rudimentary Portuguese was much different than present-day phenomena under similar circumstances". Lipski's lack of documentation makes it difficult to refute his theory, but the sociolinguistic setting described above suggests a cautious evaluation of this hypothesis. On the one hand, there are significant differences between 15 and 16th century Congo and present-day Angola (cf. section 2.6). On the other hand, Portuguese-based pidgins or creoles have never been attested in contemporary Angola[4].

### 3.2.2 The Ndongo Kingdom

The sociolinguistic setting described above refers to the first stage of Portuguese activity in the Congo. However, as the Portuguese empire expanded to India and Brazil and the sugar plantations in São Tomé prospered, Portugal's diplomatic and missionary interest in the Congo dwindled. The Portuguese had not found Prester John, they had not got access to the copper mines of the Congo, and they were not getting as many slaves as the rapid expansion of their overseas empire demanded. In turn, Afonso I of the Congo had not received the skilled workers, technicians and missionaries he had been promised by the Portuguese. Moreover, he had to deal with the constant raids of slave traders from São Tomé. Hence, by the second decade of the 16th century it was clear to both him and Manuel I of Portugal that their bilateral agreement was not meeting their expectations.

The tension in the relations between Portugal and the Congo was exacerbated by the emergence of new independent states eager to take over the slave trade, which was officially under Portuguese control. One such state was Kimbundu-speaking Ndongo, south of the Congo in what is now Angola, which was then ruled by a king called the Ngola (Iliffe 1999:185). Ndongo was known to the Portuguese at least since 1504, not only because it was the main supplier of the slaves brought to Mpinda by the envoys of Afonso I, but also because, despite D. Manuel's 1504 prohibition of travel and trade beyond the Congo River, individual Portuguese had been trading along the coast of Ndongo. Moreover, there were reports that Ndongo was rich not only in slaves but also in silver and gold (Birmingham 1966:25-9).

Portugal's first official contacts with Ndongo occurred at King Ngola's request, first in 1520 and latter on in 1561. On both occasions Ngola revealed his interest in converting to Christianity and establishing a political and commercial agreement with Portugal, but when the Portuguese ambassador arrived at his court in Kabasa, he changed his mind and took the envoys hostage.

[4] Barros (2002:35) and Mingas (2002a:45) do suggest the existence of an ongoing process of creolization in Angola, but neither of them presents linguistic data to support their hypothesis, nor do they specify whether this process is occurring throughout Angola or just in some areas, namely in the rural areas as suggested by Lipski (2005:63).

Nonetheless, in the period between the two dates, illegal slave trade from the mouth of the Kwanza River prospered and a church was built by the small population of European and African men and women from São Tomé who had settled there (Russel-Wood 1998-2000a:243). Ngola's attitude towards the Portuguese was most probably influenced by the Afro-Portuguese traders from São Tomé living at his court, who were not willing to give up their tax-free supply of slaves (Birmingham 1966).

The Ngola's attitude towards the Portuguese envoys, the political crisis in the Congo that followed the capture of Mbanza Congo by the Jaga[5] in 1569, and the shift of the Portuguese trading base from the Congo to Ndongo led the Portuguese crown to reconsider its policy on the exploration of the western coast of Africa. Paulo Dias de Novais, who had been taken hostage by the Ngola in 1561 and released in 1565 was a key factor in the implementation of this new Portuguese policy, i.e. military conquest (Amaral 2000). Therefore, according to Birmingham (1966:40), 1565 "can conveniently be taken to conclude the first stage of the history of Portuguese activity in West Central Africa".

## 3.2. The founding of Luanda: the early days of the colony of Angola (1575-1640)

The illegal slave trade that prospered from the 1530s to the 1560s developed around the region of Luanda, which was governed by a petty king that paid tribute both to the king of the Congo and the Ngola[6]. This had long become the new Portuguese trading base, but the Portuguese had no firm hold there. So, in order to secure it, King Sebastião of Portugal nominated Paulo Dias de Novais as the first governor of the new colony of Angola, named after King Ngola. The charter ceding Angola to Novais was the foundation for a renewed Portuguese expansionist policy:

> A carta régia de Novais caracterizava-se pela falta de vontade dos Portugueses em perpetuar o *status quo*. Não foi apresentada ao ngola a possibilidade de negociação. A carta régia começava por afirmar que o reino de Angola devia ser dominado e conquistado, que iria verificar-se a conversão e que os benefícios comerciais seriam para D. Sebastião e Portugal (Russel-Wood 1998-2000a:243)

Novais was not given any financial help, but he was awarded the area from the River Cuanza to the River Dande, including thirty-five more leagues south of the former, with no limits imposed on Portuguese progression towards the interior. Novais was also allowed to give land to those farmers he found deserving of it, and he had full jurisdiction over judicial, military and fiscal

---

[5] Group of unknown origin that invaded the Congo in the 16th century.

[6] In fact, this region was the supplier of the *nzinbu*, i.e. the small shells that were the official currency in the Congo (Vansina 1997:222).

affairs. In return for these benefits, the new governor was expected to implement a detailed plan of settlement and colonization. Hence, within less than two years after leaving Lisbon he was expected to train 400 soldiers and to attract skilled Portuguese workers to the colony; within six years he was to settle 100 families and ten years after his arrival Novais was expected to have built three forts and several churches (Amaral 2000:49-72).

In 1575, Novais arrived on the island of Luanda, whose population was estimated at 3,000 (Birmingham 1966:48) and in 1576 he moved with some settlers to the mainland, where they founded the city of São Paulo de Luanda. The Portuguese were received by the "principais da terra" (i.e. local chiefs) and probably by some of the 40 wealthy Portuguese who had fled the Congo on occasion of the Jaga invasions. The present writer was not able to find any references to the racial composition of the people that accompanied Novais or to the languages they spoke, but there is evidence that an interpreter was used in communications between Paulo Dias de Novais and the ambassadors of the Ngola. However, the use of this interpreter was apparently more a ritual than an actual need because Novais spoke fluent Kimbundu as a consequence of his four-year captivity at the court of the Ngola. What is interesting is that the interpreter used on this occasion was a white man said to speak Kimbundu as correctly as the ambassador himself (Amaral 2000:92), hence showing that "on the mainland it was more a matter of the Portuguese traders and adventurers becoming Africanized than of the Negros becoming Europeanized" (Boxer 1963:13).

In the case of 16th century Angola this statement was particularly true. The main reason for this was the very limited number of metropolitan Portuguese who settled there as well as the social background of those who did. In fact, the majority of these Portuguese came from other parts of the African empire, especially from São Tomé. They were mostly criminals and convicts who had entered the slave trade and had long abandoned the Portuguese language and culture[7]. In fact, their linguistic knowledge and their diplomatic connections in the interior was the main reason why these convicts and their descendants were tolerated by the Portuguese authorities in Angola until the late 19th century (Russel-Wood 1998-2000b:262).

The number of Portuguese women in the colony was even lower than in other parts of the Portuguese empire (Boxer 1977:28-35). The largest immigration of metropolitan women to Angola in this period occurred in 1595 with the arrival of 12 converts from the Casa Pia das

[7] In fact, those already living in the island of Luanda at the arrival of Paulo Dias de Novais were not comfortable with the idea of being summoned by the governor as he represented the return of the metropolitan law and taxes (Amaral 2000:89).

Convertidas (Amaral 2000:103)[8]. The shortage of white women in the colony meant that "most of the children of settlers were educated by their African mothers and the slave women in the household, so that their mother tongue was African" (Vansina 2001:269).

The reduced number of Portuguese in Angola was closely related not only to their lack of resistance to malaria and to the harsh tropical climate, but also to hunger, as a consequence of the climate's being unsuitable to their agriculture, as shown by Santos (1998:89):

> Certos surtos endémicos violentos, como os registados em 1576-1577, 1585 e 1626, e que dizimam a população europeia radicada em Angola, têm como causa imediata o flagelo da fome. A endemia de 1581, devida em grande parte à fome, segundo o P^e^. Baltasar Afonso, vitima, no espaço de oito meses, cerca de 100 portugueses e 40 angolanos do exército conquistador e, ao fim do ano, já tem imoladas duas partes dos 300 soldados brancos.
>
> Os efeitos do surto de 1595 não são menos catastróficos: mata mais de 200 europeus empenhados nas guerras do Bengo (...)

The continuous wars with the surrounding African kingdoms from 1579 until 1690 were another major constraint on the number of Portuguese settlers in Angola. In fact, the first years of the colony were characterized by friendly, peaceful relations with the Ngola. However, as the Portuguese pushed farther into the interior and approached the limits of the Ndongo and other Mbundu states, war eventually broke out (Russel-Wood 1998-2000c:128). These wars had serious consequences on the already reduced number of Portuguese settlers in Angola. It is estimated that 2,340 Portuguese, mostly soldiers, were sent to the colony from 1575 to 1592, but only 300 were still in Luanda in 1592 because 450 had been killed in the wars and the others had either died of malaria or escaped into the interior, where they acquired the African culture and languages (Santos 1998:85). The high rates of mortality among the Portuguese military meant that "the army depended largely on African slave-soldiers and the private armies of allied or conquered Mbundu chiefs" (Birmingham 1966:52-3).

Angola's total dependency on the slave trade also constrained the success of the colony's settlement policy as it distracted attention from activities such as agriculture and fishing which might have helped create conditions more conducive to the development of a colony suitable for permanent settlement. In fact, the Portuguese penetration of the interior was more motivated by the wish for slaves than opening up permanent settlement areas. A clear example of this was the construction of forts in areas around Luanda along the Kwanza River such as Massangano, Cambande and Lumbo in the 1580s. In fact, according to Klein (2002:77), "todos os postos avançados dos europeus tinham por missão fundamental manter abertos os canais comerciais

---

[8] These were orphans girls who had been living at the religious institution called *Casa Pia das Convertidas*.

com os africanos ou garantir a continuidade do comércio, não tendo o propósito de crier enclaves nacionais ou coloniais".

The founding charter further emphasized the development of a "colónia de exploração" rather than the establishment of a "colónia de povoamento" because it was financed by individual Portuguese rather than by the Portuguese crown. Hence, the interests of the latter were often overridden by those of the former, whose only purpose was to accumulate wealth and return to Portugal (Santos 1998:91).

Due to all these factors, when Paulo Dias de Novais died, in Massangano in 1589, the Portuguese presence in Angola was restricted to Luanda, a few forts in the interior along the River Kwanza and the city of Benguela, "none of which were more than two hundred miles from the coast" (Boxer 1963:39). Portuguese control of the coastal forts was secured not by the Portuguese themselves, whose presence was never numerically significant, but by an Afro-Portuguese elite, the outcome of the high degree of miscegenation between Portuguese men and African women (Russel-Wood 1992:60-1). Klein's (2002:65) statement below about the role of the Afro-Portuguese in Angola's slave trade is revealing:

> ... os portugueses instalaram colonos em Luanda e depois em Benguela, e aí nasceram, fruto desses esforços e de mercadores locais que se deslocaram para o interior, as maiores comunidades afro-portuguesas da África Ocidental. Embora os clãs de mercadores africanos controlassem a maior parte do comércio do norte de Luanda até à margem sul do Zaire, a maior parte do comércio de Luanda para sul estava nas mãos destes comerciantes afro-portugueses, pelo menos ao longo da costa e um pouco para o interior.

Bilingual in the Portuguese of their fathers and in the Bantu languages of their mothers, the Afro-Portuguese not only acted as slave hunters and protectors of the trade routes in the interior, but also occupied high positions in the public administration in Luanda and Benguela.

The social setting described above is consistent with the hypothesis that from the foundation of Luanda in 1576 until 1600, Kikongo and later Kimbundu were the most widely spoken languages in the colony. In the words of Vansina (2001:270) "while the immigrants did ensure that Portuguese remained the official language of the colony, they were not able to further increase the use of that language among the Africans in the city nor among the settler community".

## 3.3. The Dutch capture of Angola and Brazilian rule of the colony (1641-1700)

As shown in the previous section, the first fifty years of the colony of Angola were characterized by extensive wars with the neighboring African states for control of the slave trade and the silver mines which the Portuguese thought abounded in Angola. However, by the 1630s the Portuguese had managed to secure the area around Luanda, establishing peace agreements with most of the local chiefs. According to Vansina (2001:270) "the settler community in Luanda was still small, but it was wealthy and thrived" like other centers in the interior:

> Mas, também, o interior da colónia patenteia alterações significativas: aumento da população dos presídios e alastramento da mancha de povoamento. Nascerão, deste modo, outras pequenas povoações de colonizadores e Quionzo, Golungo, Bemba, Nambacalombe e Quilunda, como surgirão, ao longo dos rios Cuanza, Dande e Bengo, muitas «fazendas», com as suas casas térreas, seus terreiros e outras dependências anexas, nomeadamente capelas (Santos 1998:94).

However, the development of the colony suffered a setback in 1641, when the Dutch captured Luanda and Benguela with the help of King Garcia II of the Congo and Queen Nzinga of the Jagas. The Portuguese resisted the Dutch attacks on Luanda for three days, but they were finally forced to abandon the city, taking refuge near the River Bengo, where the Jesuits had their agricultural estates (Magalhães 1998-2000:80-1).

The Dutch ruled Angola and Benguela for seven years, until in 1648, when the Luso-Brazilian Salvador Correia de Sá e Benevides succeeded in recapturing the two cities with an expedition of two thousand men. From this date onwards Angola was ruled from Brazil rather than Portugal itself as the governors appointed by the Portuguese king came mostly from Pernambuco (Magalhães 1998-2000:81) as did the first settlers in southern Angola in the 18th century. The precise linguistic consequences of this influx of Brazilians in the colony are unknown, but it is significant that as late as the 19th century a number of travellers noted the similarity between Brazilian and Angolan Portuguese (cf. chapter 2, section 2.1.3).

One would expect that in the years following the Portuguese recovery of Luanda and Benguela, the Portuguese crown would have invested in the effective settlement of Angola. However, this was not what happened. The nomination of Brazilian governors had the main purpose of speeding up the flow of slaves to Brazil and the Spanish American colonies. Hence, the thirty years that followed the defeat of the Dutch were characterized by consecutive wars against the Mbundu states of Matamba, Kissama and Kassange, which had not only helped the Dutch but now also imposed severe constraints on the access of the Portuguese to the potential slaves in

the interior. Securing extensive supplies of slaves was, once again, the main concern of the Portuguese crown in Angola.

As far as the sociolinguistic setting is concerned, Vansina (2001) argues that in the interior the use of Portuguese as a lingua franca expanded among chiefs and traders essentially because "reading and writing was useful ... to establish political rights, to secure inheritances by written wills, and to claim or object to claims concerning trading matters" (*ibid.* 272). However, it is likely that this expansion was also due to the fact that most of the few Portuguese whom the crown was able to draft to go to Angola were sent to the interior, where they were most needed to secure the *presídios* (i.e. forts) at a time of ongoing warfare with the Mbundu states. In fact, according to Boxer (1965:133):

> After the recovery of Luanda in 1648, it became customary to send out drafts of recruits for the garrison with the incoming governor every three years. The drafts were mainly raised from the overpopulated islands of Madeira and the Azores, particularly the former.

A significant number of these draftees came also from Pernambuco and Ceará in Brazil, as they were accustomed to the tropics and were more likely to survive in Angola (Boxer 1965:134). Regardless of where they came from, these draftees were mostly composed of *degredados* or convicts.

In contrast to this expansion of Portuguese as a lingua franca in the immediate vicinity of the forts in the interior, Kimbundu was the most widespread language in nearly all households in 17th century Luanda and in the daily life of the city. According to Vansina (2001:271), the factor that contributed the most to this was the fact that, owing to the high rate of mortality among European settlers and the low influx of women into the colony, the large group of Afro-Portuguese "began to occupy more and more positions in the army and the local administration". Hence:

> As a result, Kimbundu, especially from the 1650's onwards, gradually came to be as essential in administration, the army, the church as it was for inland commerce, despite the official status of Portuguese and despite the fact that high officials or bishops sent out from Lisbon only knew Portuguese (*ibid*).

Another important factor for the consolidation of Kimbundu's position as the most widely spoken language in Angola was the fact that the majority of the slaves arriving in Luanda in the 17th century were Kimbundu speakers, as they were mainly drawn from the crumbling Mbundu states of Matamba, Kissama and Kassange. Consequently, Kimbundu was also the official language of missionary work, as shown by the fact that "slaves exported from Luanda, whatever their origins, learned some Kimbundu and were baptized in this language before their

embarkation" (Vansina 2001:273). The role of Kimbundu in missionary work is also shown by the publication of several catechisms in this language. These publications were of great importance in the standardization of the language. The exception that proves the rule appears to have been a closely-knit group of black, Portuguese-speaking, Roman Catholic families known in the history of Angola as the Old Creoles of Luanda (Birmingham 2002:148)[9].

## 3.4. Angola in the 18th century

### 3.4.1. The development of Luanda

From 1683 to 1730 there was peace in Angola (Marques 1995:435). However, the expansionist policy of the colony still focused on the slave trade rather than the effective settlement of the territory. With the exception of the few forts along the Rivers Kwanza and Cunene, the effective Portuguese hold on the colony was restricted to Luanda.

In fact, with the intensification of the slave trade throughout the 18th century, Luanda was all that mattered. The city, whose population is estimated at 7204 inhabitants in 1796, was divided into two main parts: the center and the periphery. The former was also divided into two parts: (i) the *Cidade Alta*, which was the political, military and religious center of the city, and (ii) the *Cidade Baixa* or commercial zone (Venâncio 1996:32).

According to Venâncio (*ibid.* 36), in 1793 the population of the *Cidade Alta* included 210 Europeans, 114 Euro-Africans, 96 free Africans and 848 slaves. Only three years later, in 1796, the population was distributed as follows: 85 Europeans, 109 Euro-Africans and 708 Africans. Venâncio's account is interesting in two more ways. First, the use of terms *Europeans* and *Euro-Africans* as opposed to *Portuguese* and *Afro-Portuguese* suggests the presence of Europeans other than Portuguese. In fact, Venâncio acknowledges that the term European refers to the whites "com origem próxima e remota naquele continente" (*ibid.* 47), and includes Portuguese whites, Angolan whites and Brazilian whites. Second, his data show that only a small part of the existing social groups in Luanda lived in the *Cidade Alta*, i.e. 15,1% of whites, 8,6% of Afro-Portuguese and 13,1% of Africans[10], suggesting a greater concentration of all social classes on the periphery.

---

[9] Notice that the meaning of the word creole is not that of contemporary contact linguistics. Rather, the word is used here to denote the Portuguese cultural heritage of these black families.

[10] The majority of this African population were women who worked as *mocambas* 'domestic slaves' (Venâncio 1996:46). In fact, a census of the population of Angola in 1777-1778 showed that there were twice as women as there were men (Iliffe 1999:191).

In fact, the periphery, which included most of the poor houses of the Africans who worked in the *Cidade Alta* and the yards where the slaves brought from the interior awaited shipment to Brazil and other parts of the Americas[11], had a larger representation of all the social groups in the city. Hence, the ethnic composition of this part of Luanda in 1796 was as follows: 196 Europeans (i.e. 34,5% of them), 470 Afro-Portuguese (i.e. 37,7%) and 1810 Africans (i.e. 33,6%).

The exact percentage of Africans living in Luanda was constrained by the dynamics of the slave trade. For example, from 1781 onwards their number decreased, to a large extent due to the reduction in the number of slaves brought to Luanda (*ibid.* 47). Nonetheless, Africans always made up the majority of the city's population. The languages spoken by this group are likely to have been related to the area of the city they occupied. In fact, those living in the *Cidade Alta* are likely to have spoken Portuguese, despite the fact that their mother tongue was probably African. Those living on the periphery were most likely to have spoken Kimbundu[12].

While the Afro-Portuguese, i.e. mixed-race, ranked third in the population of the *Cidade Alta* they were in fact the second largest social group in Luanda. Their reduced numbers in the wealthier parts of the city is probably due to the fact that most of them were *degredados* who worked as slave traders and hence stayed with their property on the periphery. Nonetheless, regardless of the area of the city they lived in, they also occupied the majority of public offices in Luanda (*ibid.* 50) and their mother tongue was mostly Kimbundu (*ibid.* 53), despite their bilingualism in Portuguese.

---

[11] According to Klein (2002:129) "a maioria dos escravos passava um mínimo de seis meses a um ano desde a captura ao embarque nos navios europeus, sendo que o tempo passado na costa à espera do embarque era em média de três meses". It is not likely that slaves acquired any proficiency in Portuguese during their stay in Luanda, not only because three months was a too short period, but above all because most of the time waiting for embarkation was spent in the interior, where traders kept slaves for as long as it took to Europeans raise their bids (*ibid.* 90). During this period, slaves were used as carriers or farmers and are likely to have acquired at least basic proficiency in the African languages spoken by slaves working in those communities in the interior. This hypothesis is supported by Klein's following statement (*ibid.* 155 – my emphasis):

> Assim, um dos poucos relatos de testemunha presencial da captura e transporte de escravo, feito pelo Igbo Olaudah Equiano, feito escravo em 1750, ressalva com evidência que este rapaz passou pelas mãos de vários vendedores antes de chegar à costa. Além disso, passou determinada altura um mês inteiro a viver uma vida relativamente normal, tendo mesmo trabalhado como aprendiz de ourives. *Passou metade do tempo com povos que falavam línguas parecidas com a sua, várias das quais aprendeu, e participou na economia local.*

[12] In the 18th century most slaves came from the region of the Lunda, in the northeast (Iliffe 1999:199), but they were likely to speak Kimbundu also, not only because this was the mother tongue of the majority of Africans and Afro-Portuguese who lived with them in the periphery of Luanda but also because they were brought from the interior to Luanda by Kimbundu-speaking Africans who worked as intermediaries between the Portuguese and the Lunda chiefs. Hence, it is unlikely that a Portuguese-based pidgin or creole was spoken by these slaves as they could use Kimbundu, which was the *lingua franca* used in slave trade in the interior.

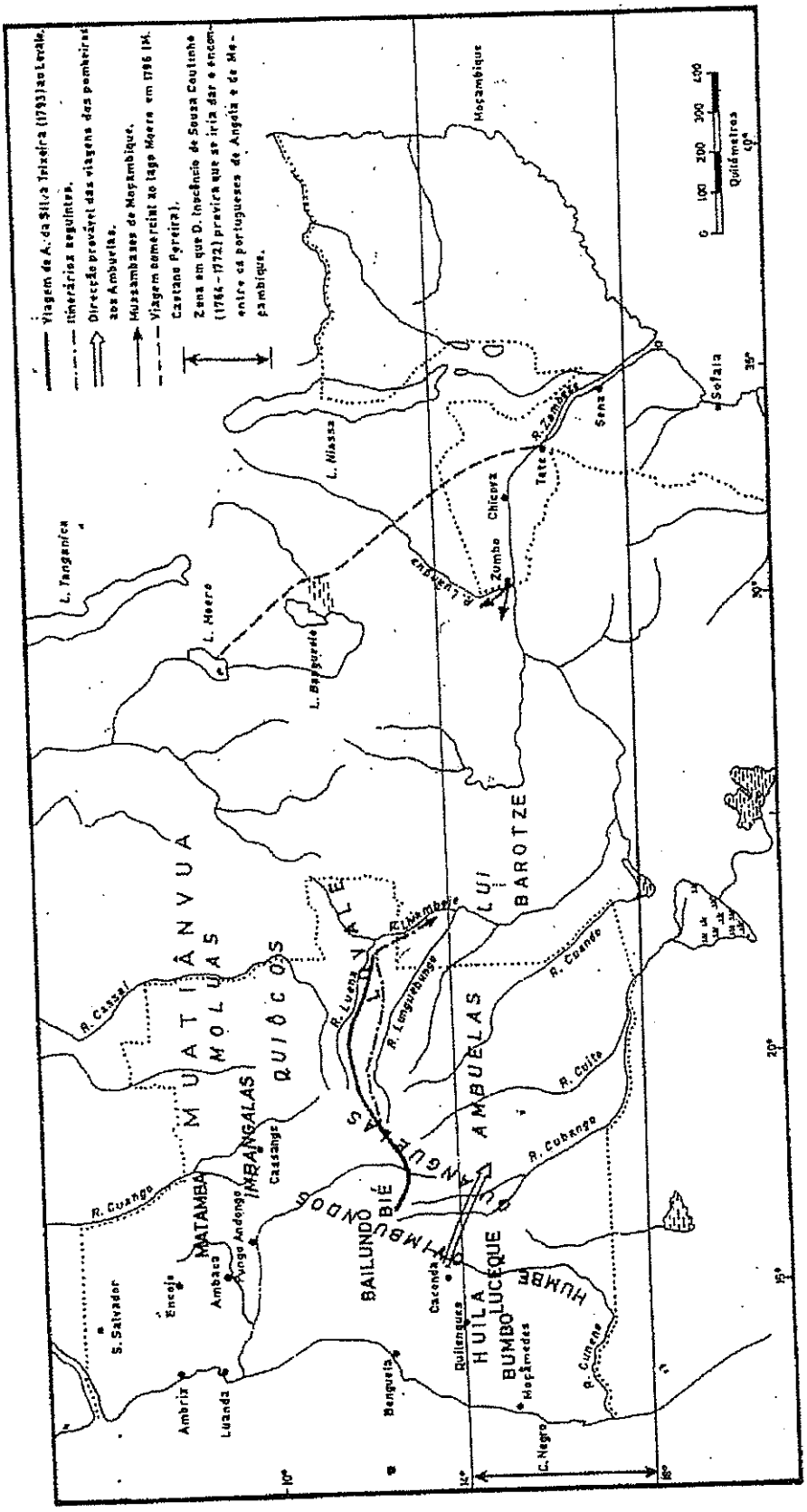
The white population in Luanda was still small but it was much more representative than in previous periods in the city's history. The growth of the white population in Angola led to the emergence of the first conflicts between the Portuguese and the Afro-Portuguese, mostly as a consequence of the nationalist policies enforced by the governors appointed by the Marquês de Pombal. In fact, throughout the 18th century, the Portuguese resisted the growing Africanization of the Afro-Portuguese elite in Angola. For example, in 1760, the Marquês de Pombal ordered the closing of the Jesuit school, as he considered them responsible for the diffusion of Kimbundu rather than Portuguese. This measure was reinforced by Sousa Coutinho's decree in 1765 compelling all heads of households to enforce the use of Portuguese in the home, in the education of their children and in contacts with slaves. However, while better represented, the Portuguese were still too few to enforce the imposition of their language. The situation remained unaltered until the mid-19th century. Only after this period did Portuguese gradually become more widely spoken in the colony (Vansina 2001:274-5).

It is unknown whether the sociolinguistic composition in the settlements in the interior was similar to those in Luanda, but in the latter native speakers of Kimbundu clearly outnumbered those of Portuguese.

### 3.4.2. The exploration of the south

Another important feature of 18th century Angola was the attempt to explore the southern part of the colony. In the late 16th century, when Paulo Dias de Novais founded Luanda and the colony of Angola, the River Kwanza had been a key factor in the Portuguese penetration of the interior. In the late 17th century and early 18th century the Cunene River served the same purpose in their exploration of the Benguela region (cf. Map 5). In fact, this was the main area of intervention of the Portuguese during this period, largely due to the need to find further supplies of slaves but also due to the Portuguese wish to find a direct route to connect Angola and Mozambique by land (Santos 1988:135-41).

**Map 5: Portuguese exploration journeys in Central Africa in late 18th century (Santos 1988:167)**



The Benguela region had been known to the Portuguese since the first journey of Diogo Cão to the Congo and part of it was included in the founding charter given to Paulo Dias de Novais. In fact, the exploration of Benguela's hinterland began in 1578 when Benguela-a-Velha was founded. In 1592, Domingos de Abreu Pinto, in his detailed report on the colony of Angola, noted the importance of the permanent settlement of that area as a means of connecting Angola and Mozambique by land and suggested that Filipe I assign a governor to the region. Of course, Pinto's recommendation was based on an assumption that the Cunene was connected to Zambezi River, which was disproved in the 19th century. Still, in 1617 Manuel Cerveira Pereira was appointed governor of Benguela and ordered the exploration of its hinterland. The harsh tropical climate, the lack of food, the mutiny of the *degredados* who accompanied him and the threat of a Dutch attack led him to abandon the enterprise. Nonetheless, in 1638 the settlement of Caconda was founded and in 1694 a fort was built on the beach in Benguela (Santos 1988:139) as a consequence of the revived interest in the River Cunene in 1664 (Padrão 1998:30).

These explorations of the hinterland of Benguela did not result in any consequent policy of settlement. In fact, were it not for the fact that from the late 18th century onwards it received most of the slaves that were shipped from Angola to Brazil, the region would have remained in almost total abandonment until the early 20th century. Nonetheless, these exploratory journeys provided the Portuguese with important geographical information and the will to develop that region. As in most of Angola, the first attempts at settlement occurred only in the mid-nineteenth century, when the specter of the Berlin Treaty already loomed over Portuguese interests in the colony and Brazil had already become independent.

## 3.5. The 19th century and the start of Portugal's African Empire: the case of Angola

The key event in the history of the Portuguese empire in the 19th century was the independence of Brazil in 1822. This was not only because this event represented the end of the second Portuguese overseas empire, but above all because it forced the Portuguese crown to focus on Africa for the first time since the beginning of overseas expansion[13]. In fact, in the years immediately following the independence of Brazil there was a widespread perception that unless the effective Portuguese holding in Africa were reinforced and expanded and commercial

[13] Despite the long-held myth in Portugal that there was only one Portuguese overseas empire and that it extended for a period of five hundred years, contemporary Portuguese historians tend to refer to the existence of three distinct empires: the Asian empire (c. 16th century), the Brazilian empire (c. 16th to the 19th century) and the African empire (c. late 19th century to 1975) (Alexandre 2000:11).

relations with the African colonies were increased, assets as important as Angola could follow Brazil's path to independence (Alexandre & Dias 1998:23-ff).

The first step in bringing the metropolis closer to the African colonies was to put the latter on an equal administrative footing with the former. Hence, the liberal constitutions of 1822 and 1826 established that the colonies were to become provinces of Portugal and as such were entitled to representation in the parliament. The new provinces were also divided into districts, which in turn were divided into *concelhos*, and their laws and public offices were those of the metropolis (cf. Dias 1991:15-32). Moreover, there was a clear intention to enforce effective settlement. Concerning Angola, the plan was to secure the north by building several establishments along the coast from Ambriz to the Congo, to reinforce the connection between Luanda and Benguela, and to explore the coast and the hinterland of the latter. However, these intentions had no effect on the settlement of the colony as a whole:

> Por volta de 1830 ou 1840, a ocupação efectiva de Angola reduzia-se a uma estreita faixa costeira de 80 a 130 km de largura média, sem limites definidos para o interior. No Norte, o território ocupado alargava-se um tanto, abrangendo as duas margens do Zaire até à actual Matadi, 171 km distante da foz, onde a navegabilidade do rio se detém nas quedas de água Livingstone. No Sul o sueste de Benguela de hoje, a soberania portuguesa alcançava Caconda, uns 240 km a oriente do mar. A própria linha costeira só se conhecia com pormenor até Benguela antes da expedição naval das décadas de 1830 e 1840 que a traçou de forma científica até à Baía dos Tigres, 17° de latitude Sul. (Marques 1998:140-1)

Moreover, the population in the centers administered by the Portuguese was extremely limited (i.e. 250,000 to 300,000), especially the white population (i.e. 1000 to 2000 in total). The bulk of the population in the Portuguese settlements was made up of Afro-Portuguese settlers, officers and merchants. In short, in the early 19th century

> A presença portuguesa em Angola devia a sua sobrevivência à imbricação complexa dos interesses mútuos do governo colonial, das famílias Afro-Portuguesas do litoral e interior da colónia e dos titulares políticos dos «Estados» africanos autónomos do sertão. (Dias 1998:366)

In fact, it was not until the 1850s that new settlements started in Angola, mostly due to the action of the Marquês de Sá da Bandeira, whose policy emphasized territorial expansion, the effective end of the slave trade and the development of a plantation economy (Alexandre 1998:84)[14]. The enforcement of the plan implied large numbers of Portuguese settlers, but this was only possible in Moçâmedes, Porto Alexandre and Baía dos Tigres due to a the arrival of a

[14] In fact, according to Dias (1998:357), until the 1850s the effective administration of the Portuguese in Angola was restricted to nine *presídios* (Ambaca, Duque de Bragança, Cambambe, Massangano, Muxima, Novo Redondo, Pungo Andongo, S. José de Encoje, Benguela e Caconda) and thirteen districts (Barro do Bengo, Barra do Dande, Barra de Calumbo, Dande, Icolo e Bengo, Zenza e Quilengues, Dembos e Golungo, Dombe Grande, Bailundo, Bié, Huambo, Quilengues de Benguela).

growing number of settlers from Pernambuco and the Algarve (Padrão 1998). The present writer was not able to find any information regarding the ethnic origin, sex or age of these settlers, but it is likely that the ratio between blacks and whites was more balanced than in Luanda or in those countries where Portuguese-based creoles developed. While there are no completely reliable demographic data for this period and this region, it is useful to present the estimate quoted by Pélissier (1997:192) with regard to the demographics in the district of Moçâmedes:

**Table 2: Demographics in the District of Moçâmedes (1854 and 1864) (Pélissier 1997:192)**

| | 1854 | 1864 |
|---|---|---|
| **Brancos** | 256 | 1124 + 307 militares |
| **Mestiços** | 29 | 127 |
| **Negros livres** | 59 | 862 |
| **Libertos** | 135 | 275 |
| **Escravos** | 481 | 2073 |

As far as language is concerned, it is important to notice that there are references in the literature to the development of a mixed language among the population working for the Portuguese in Moçâmedes:

> ... a sociedade Africana da zona costeira de Moçâmedes formou-se a partir de uma população escrava, composta de libertos, do tráfico ilícito, cativos de guerras coloniais, e um número crescente de vítimas do tráfico entre as feitorias portuguesas na costa ao sul de Luanda e as sociedades autónomas do interior. Oriundos na maior parte, de diversas regiões a norte e leste do Cuanza, este *grupo heterogéneo* de pessoas desenvolver uma *identidade quimbari distinta, baseada na língua quimbundu* e numa *cultura sincrética, afro-portuguesa. Tornou-se assim culturalmente distinta de outros trabalhadores contratados mais tarde*, oriundos sobretudo do planalto central, os quais mantiveram um pé nas suas comunidades de origem, embora alguns tivessem sido absorvidos na comunidade *quimbari* (Dias 1998:446 – my emphasis).

What is interesting about Olumbali, the language spoken by the *Quimbari*, is that it shows that it was the African languages and not Portuguese that were more likely to undergo significant restructuring in Angola. In fact, there are several references in the literature to African mixed-languages in Angola, especially in borderline areas of the Bakongo, Ambundu and Kimbundu territories (cf. discussion in chapter 2, section 2.1.4).

Huíla was another area that received particular attention from the Portuguese crown during the mandate of Sá da Bandeira as president of the Concelho Ultramarino (i.e. 1856-1859)[15]. In 1857 he created the Kruss colony, a military settlement made up of twelve interns from the *Casa Pia* and twenty-nine Germans. As far as Huíla is concerned, it is interesting to notice Padrão's (1998:137) comment concerning the ethnic composition of its settler community:

> Um dos aspectos mais interessantes da colonização da Huíla é o do papel que muitos mestiços e autóctones desempenharam na consolidação da soberania de Portugal.

These experiences of effective settlement were not only innovative in comparison to previous experiences, but they were also profitable at first and contributed immensely to the increment of cotton and coffee production in the colony. An important factor was that the settlers in some of these communities included experienced Brazilian technicians and *feitores* (i.e. foremen). However, their success, especially in the interior, was short-lived. First, as there was almost no infrastructure for internal communication and travel, it was extremely expensive to transport produce from the interior to the coast, especially after Sá da Bandeira's prohibition on the hiring of African carriers. Second, while the Portuguese crown supported the creation of these communities politically, its financial support of these enterprises was very small.

However, the most important factor in the decline of the communities in the south was the fact that, in the 1860s, after Sá da Bandeira left the presidency of the *Concelho Ultramarino* following strong opposition to his policy from the governors in Angola and politicians in the metropolis, the official Portuguese attitude towards these communities changed once again. In fact, all ongoing campaigns for exploration and settlement were suspended, not only in the interior but also on the coast, and attention turned again to stimulating the slave trade as this was perceived as the only way to not only increase the profits of the colony but also to free Africans from their inherent savagery (Alexandre 1998:84). If to this we add the violent insurrections of several African leaders that started in the south in the 1860s, we have all the ingredients to depict Africa and Angola in particular as a very dangerous place to invest in. This, of course, had damaging consequences on the number of white settlers in the regions that had attracted such a considerable quantity of them in the 1850s (cf. Pélissier 1997:200).

Nonetheless, the 1870s, 1880s and 1890s brought about a renewed interest in Angola, not only because Portugal had achieved political and economic stability and the profits from the rubber trade in Angola were significant, but above all because other European colonial powers showed

[15] The *Concelho Ultramarino* was first created by John IV of Portugal in 1643. It was abolished in 1821 and recovered again in 1851. Throughout the century it was always the institution responsible administering everything related to the overseas territories (Dias 1991:22-32).

a growing interest in Africa in general and in the Portuguese possessions there in particular (Alexandre 2000:17). In fact, the most important event in the 1880s in this respect was the Conference of Berlin in 1884, not only because it defined Angola's northern and south-eastern borders but above all because it triggered an extreme Portuguese nationalist feeling which led to the depiction of the African empire in general, and Angola in particular, as a symbol of Portuguese identity which had to be maintained and protected against all attacks.

Hence, the last three decades of the 19th century were characterized by the proliferation of military campaigns and the establishment of new settlements, especially in the Huíla region. With the exception of the community of São Januário and Chibia, which were founded by Boer refugees from South Africa, the two most important settlements founded in Huíla were Sá da Bandeira and Lubango, both founded in 1885 by *degredados* from the overpopulated island of Madeira. As late as the mid-twentieth century, Sá da Bandeira and Lubango were to remain the only places in Angola where Europeans outnumbered Africans.

This period was also characterized by the organization of geo-military expeditions to regions in the interior where conflicts were likely to erupt when the Portuguese tried to establish a firmer hold there. The most important of those areas was the Lunda, which from the 18th century onwards had provided Luanda with the majority of the slaves shipped to Brazil and the Americas (Dias 1998:335). It was also during the last three decades of the 19th century that the first investments were made in developing an embryonic net of internal communications, such the railways connecting Luanda with the interior in 1887 (Alexandre 1998:114).

The 19th century also brought about significant sociolinguistic changes in Luanda as the white metropolitan population in the city doubled (Vansina 2001:276). The influx of white settlers started right after Brazil's independence from Portugal, as Portuguese military were sent to Luanda to ensure that the faction of the Afro-Portuguese elite that supported the integration of Angola under Brazilian rule, i.e. to become a colony of Brazil, would not succeed (Alexandre 1998:25-6). In the following years and throughout the 19th century the white population in Luanda increased as a consequence of the continuing arrival of settlers to the communities in the south, who often either stayed in Luanda or returned there when the aforementioned settlements failed to prosper. While most of these settlers were convicts and criminals, they also included skilled workers and the sons of the leading families in Luanda who had been sent to Brazil to study. Moreover, the growing presence of whites in Luanda brought about some investment in the creation of facilities for recreation (e.g. the theatre created in 1839) and education (e.g. the public school for boys opened in 1835 and another one for girls founded ten

years later) as well as the development of an embryonic social life in which women, especially white women, started to have a role (Dias 1998:508-512; Vansina 2001:276).

According to Vansina (2001:277) all these factors led "more men from the local elite to acquire a more thorough knowledge of Portuguese and to use the language more" and consequently "the Luso-African community eventually became completely bilingual". Vansina's hypothesis is corroborated by the American missionary Heli Chatelain, who states in the introduction to his compilation of Kimbundu folk-tales that he "was compelled to master Portuguese, which in Angola is indispensable for dealing with the educated classes, and is always of utility in the intercourse with the common natives" (Chatelain 1894:v). The bilingualism of the Afro-Portuguese elite in Angola is also attested by the publication of a bilingual weekly newspaper in Luanda in 1882 (Marques 1998:161)[16].

However, despite the increasing influence of metropolitan Portuguese, it is likely that the changes in the sociolinguistic setting were more significant in Luanda rather than in the interior of the district[17]. The following remarks by Heli Chatelain cast light on the relevance of Kimbundu as a lingua franca in the district of Luanda for people of all colors and classes:

> ... the political and commercial importance of the Loanda district, where Ki-mbundu is the vernacular, the number and partial civilization of the inhabitants, the vast extra-territorial use of the language – in the coast-belt, wherever there are to be found Portuguese traders, troops, or authorities, and eastwards as far as the Lualaba, wherever the ubiquitous Ambaquista (native of Mbaka) has penetrated – fully warranted the founding of a Kimbundu literature ... (Chatelain 1894:vi).
>
> The autodidactic and practical Ambaquistas of the interior have begun to perceive the superiority, for purposes of private correspondence, of their own tongue to the Portuguese, - to them what Latin is to the Lusitanian peasant; finally, indications are not wanting that the Portuguese authorities, civil and ecclesiastic, are becoming awake of the importance of a general language like the Ki-mbundu as a link between the official speech and the multitudinous Bantu dialects of their vast province of Angola. (*ibid.* viii)

[16] The last decades of the 19th century witnessed the proliferation of several newspapers and magazines in Angola. For a list of the most important of these publication in the colony during this period cf. Dias (1998:529-31).

[17] Chatelain (1888-1889:XII) described the limits of the district of Luanda, and hence that of the area where Kimbundu was spoken natively by the majority of the population, as follows:

> A área em que se falla o kimbundu e aproximadamente igual á do actual Districto de Loanda, e comprehende portanto os concelhos de Loanda, Calumbo, Muxima, Massangano, Cambambe (Dondo), Pungo Andongo, Malange, Duque de Bragança, Ambaca, Cazengo, Golungo Alto, Zenza do Golungo, Alto Dande, Encoje, Ambriz, Barra do Dande, Barra do Bengo e Icolo e Bengo.

> O dialecto de Loanda é fallado no concelho d'este nome, isto é na capital e nos seus arredores immediatos, em toda a província pelas pessoas oriundas da metrópole, que accompanham ou precedem os brancos como criados, officiais, caixeiros, funccionarios publicos, traficantes ou aventureiros independentes. (Chatelain 1888-1889:XIV)

In fact, even in Luanda and the surrounding areas, Portuguese was not able to supersede Kimbundu as the most widely spoken language. The two languages continued to co-exist in what appears to have been the beginning of a diglossic situation that has survived in Angola as a whole to the present. The members of the Afro-Portuguese elite or *filhos do país,* such as Manuel Francina and Alfredo Trony, who invested in the study and dissemination of the Kimbundu language, folklore and proverbs were a determining factor in this (Dias 1998:518).

The contact between Kimbundu and Portuguese in the city and its vicinities is likely to have resulted in mutual interference. The references by Schuchardt (1888:230) to the *mestiço* language of Luanda and by Valdman (1978:22) to the *pequeno português* have been interpreted by linguists as evidence for the development of a pidginized or creolized variety of Portuguese in Angola from which the present vernacular may have sprung. However, Schuchardt's definition of *mestiço,* for example, is not consistent with this hypothesis (cf. discussion of this topic in chapter 2, section 2.1.2). If anything, *mestiço* is likely to refer to the variety of Kimbundu spoken in Luanda, which, according to Chatelain (1894:v), is "needlessly mixed with Portuguese elements" and "offers poor material for the study of the genuine Ki-mbundu".

In fact, as an appendix to his *Folk-Tales of Angola* (1894), Chatelain provides a list of 629 notes, about 90 of which refer to Portuguese loans in Kimbundu. It is interesting to notice that these loans are not only lexical (e.g. *palaia* < EP: praia 'beach', *kololo* < EP: corredor 'corridor') but also grammatical (e.g. *tandu* < EP: tanto 'so much', *poji* < EP: pois, *loko* < EP: logo). The aforementioned items refer to individual words, which have undergone significant phonetic changes. However, Chatelain (1894) also provides examples of Portuguese expressions that are used in Kimbundu as single words, e.g. *maseka* (from EP *ama seca* 'nanny'), *azalma* (from EP *às armas* 'to the weapons'), *buajtadi* (from EP *boas tardes* 'good afternoon'), *andaxi* (from EP "ainda assim" 'even still') and *dikue* (from EP *do que* 'of which').

In the light of the limited existing data, which include not only lexical items but also conjunctions and adverbial particles, one can tentatively hypothesize that the significant linguistic traits that distinguished the variety of Kimbundu spoken in Luanda from those spoken in the interior (e.g. Kimbundu spoken in Ambaca) were the result of what Thomason &

Kaufman (1988:74-76) define as slight structural borrowing[18]. The sociolinguistic setting in Angola at the time is consistent with this hypothesis. First, the borrowing process was initiated by the speakers of the borrowing population. Second, the native language of the borrowing speakers was maintained, despite their bilingualism in Portuguese. Third, while there was an increased contact with Portuguese, this was still reduced to specific contexts, as the majority of the population used mostly Kimbundu both in their daily life and in their correspondence.

It is likely that this structural borrowing continued well into the early 20th century. In fact, judging by the words of Tavares (1915:vii-viii) quoted in chapter 2 (i.e. section 2.1.4), it is likely that as the number of native speakers of Portuguese in Angola increased in the early 20th century, this structural borrowing actually intensified and was closer to level 4 or 5 of Thomason & Kaufman's five-level borrowing scale, i.e. moderate and heavy structural borrowing, respectively[19]. Similar examples of such structural borrowing can be found in the fifty tales in Cuanhama, Nhaneca and Humbe compiled in Estermann (1971). Examples include *masi* (from EP *mas* 'but'), *osaku* (from EP *saco* 'bag'), *para* (from EP *para* 'to'), *onombolo* (from EP *bolo* 'bread').

## 3.6. The 20th century

### 3.6.1. The pacification wars, the Republic and the early years of the *Estado Novo* (1900 - 1930)

While the Portuguese achieved an important expansion of their area of influence in Angola by means of military campaigns and the foundation of several agricultural settlements in the south, the effective settlement of most of the areas controlled by the Portuguese in the colony was not possible until the first three decades of the 20th century. The main reasons for the delay in the effective settlement of the territory before this date were the secondary role that Portugal played in the world economy and the fact that in the last three decades of the 19th century the Portuguese had to overcome the opposition of not only several African peoples but also that of

[18] Thomason & Kaufman (1988:37) define *borrowing* as "the *incorporation* of foreign features into a group's native language by speakers of that language: the native language is retained but is changed by the addition of the incorporated features".

[19] Tavares (1915) refers to linguistic phenomena in Kimbundu that indicate significant restructuring, e.g the tendency to use abbreviated forms of agreement, the acquisition of whole new sets of grammatical words and the substitution of native words by the assimilated forms of their Portuguese equivalents. Thomason & Kaufman (1988:74-6) identify extensive word order changes and borrowed inflectional affixes added to native words as indicators of moderate structural borrowing (level 4), whereas changes in word structure rules and addition of concord rules indicate heavy structural borrowing (level 5).

the Afro-Portuguese elite, who opposed the growing discrimination against them in favor of the newly arrived metropolitan whites[20] (Freudenthal 2001:261-ff). Hence, in 1900 Portugal had direct rule over only 10% of the Angolan territory (*ibid.* 291).

The first obstacle Portugal had to overcome in Angola was the violent resistance of the important African states that, for centuries, had maintained commercial relations with the Portuguese but always refused to allow any interference in their internal political affairs, e.g. Lunda and the peoples of the central plateau (*ibid.* 272). Hence, from the 1890s to 1920 the main focus of the Portuguese in Angola was in the "pacification" of the territory, or in other words, in its military occupation. Hence, the pacification wars were the main characteristic of this period. The most important ones were those in Bié (1890-1904), Humbe (1891-1915), Ovambo (1904-1915), Seles and Amboim (1917-1918) and Congo and Dembos (1913-1919).

The second major obstacle to Portuguese control of the territory was the ineffectiveness of its administration. Hence, the newly founded Portuguese Republic enforced a series of reforms that aimed to set the colony free from the excessively centralized administration that had prevailed previously[21]. In 1911 the Ministry of the Colonies was created and in 1914 and 1917, the overseas territories were given administrative and economic autonomy by means of the *Leis Orgânicas* and the *Cartas Orgânicas*, respectively. In order to further enforce decentralization, Angola was submitted to the system of the *Altos-Comissariados* in 1920, and in 1921 Norton de Matos was appointed *Alto-Comissário* of Angola (Dias 1991:31-32; Marques 2001:22-24)[22]. However, this decentralization policy was reversed after 1926, with the start of the *Estado Novo* (Marques 2001:25).

As far as the policy towards Africans was concerned, the first two decades of the 20th century continued the segregationist policy that had been started in the last three decades of the 19th century, establishing a clear distinction between the political and legal rights of the Portuguese as opposed to those of the *indígenas*. According to the *Carta Orgânica* of Angola (1917), in theory all the *indígenas* were granted the same civil and political rights as any white Portuguese citizen. The criteria laid down by the document were as follow:

[20] For a detailed analysis of the Afro-Portuguese opposition in Angola during the last decades of the 19th century see Dias (1998:519-542).

[21] This administrative centralization had not only had damaging economic consequences, but above all it had also been one the main reason for the growing tension between Portugal and the *filhos da terra* (i.e. Afro-Portuguese). Hence, administrative reform was needed to ensure a better economic development of the colony but also a more stable political context.

[22] Norton de Matos had already been governor of Angola from 1912 to 1915.

> ... todos os indivíduos de cor, naturais da província de Angola e de maior idade, poderão entrar no pleno uso dos direitos civis e políticos inerentes aos cidadãos portugueses, quando satisfaçam às seguintes condições: 1.° – saber ler e escrever a língua portuguesa; 2.° – possuir os meios necessários à sua subsistência e à das suas famílias; 3.° – ter bom comportamento, atestado pela autoridade administrativa da área em que reside; 4.° – diferenciar-se pelos seus usos e costumes do usual da sua raça. (quoted in Marques 2001:26)

In the first years of the *Estado Novo* the same policy was followed and established in the *Estatuto político, civil e criminal dos indígenas de Angola e Moçambique* (1926) and the *Códido de trabalho dos indígenas das colónias portuguesas de África* (1928). Nonetheless, it is clear that color had become a more important issue than in 1917, as shown in the definition of *indígena* in *Estatuto* of 1926: "... indivíduos de raça negra *ou* dela descendentes que, pela sua ilustração e costumes, se não distingam do comum daquela raça" (quoted in Marques 2001:26 – my emphasis). In the words of Freudenthal (2001:306):

> De facto, a legislação produzida entre 1926 e 1930 iria implantar os pilares da política colonial do Estado Novo em África, através de uma assimilação selectiva controlada, e eivada de preconceitos e de práticas racistas.

The ideology underlying the Portuguese policy towards the Africans in Angola in the two first decades of the 20th century and in the first years of the *Estado Novo* was a mixture of social Darwinism and condescending racism. The favourable outcome of the Conference of Berlin, which was against the interests of the more powerful European countries, and the Portuguese victory in the pacification wars had contributed to further exacerbate the radical nationalist feeling of the 1890s. This nationalism was grounded on the idea that Portugal had a colonial vocation and its people were superior to the savage African mobs, which of course was based on their assumption of a glorious past and on pride in the territorial extension of the empire. Hence, protecting and consolidating the empire was a matter of national identity (Paulo 2001:86). It was also a matter of philanthropy as the Portuguese argued that it was their moral obligation to rescue the poor Africans from savagery by bringing them into contact with civilization.

The diffusion of this ideology by means of a colonial education in Portugal from the very beginning of the 20th century was an important means to overcome three major obstacles the Portuguese had to overcome in Angola, i.e. attracting white Portuguese settlers to the colony, rethinking the economic and administrative system in the colonies and improving the scientific knowledge of those territories (Paulo 2001:31). Hence, from 1901, when the First Colonial Congress took place (1930), a powerful colonial propaganda machine was put to work:

> Na verdade tratava-se da educação colonial da metrópole, visando a socialização da ideologia colonial, a formação especializada de pessoal técnico-administrativo e a constituição de núcleos de preparação de futuros investigadores em "ciências da colonização" (*ibid*).

However, until the 1920s the effects of this propaganda in attracting more white settlers to Angola were insignificant, despite the fact that, in comparison with the previous century, the number of such settlers increased substantially. According to Freudenthal's (2001:309) analysis of the existing sources, the total population of the colony from 1897 to 1930 can be estimated as shown in table 3 below:

**Table 3. Population of the colony of Angola (1897-1930) (Frendenthal 2001:309)**

| Years | Blacks | *Mestiços* | Whites | Total |
|---|---|---|---|---|
| 1897 | 99,3% | – | 0,7% | 848 963 |
| 1898 | 98,3% | – | 1,7% | 672 082 |
| 1899 | 98,8% | – | 1,2% | 966 501 |
| 1900 | 99,7% | 0,06% | 0,2% | 4 789 946 |
| 1913 | 99,5% | 0,13% | 0,3% | 4 520 100 |
| 1920 | 99,3% | 0,14% | 0,5% | 4 278 200 |
| 1922 | 98,9% | 0,23% | 0,8% | 3 461 100 |
| 1924 | 98,6% | 0,28% | 0,3% | 3 445 902 |
| 1925 | 2 438 411 | – | – | – |
| 1927 | 98,2% | 0,4% | 1,4% | 2 989 308 |
| 1928 | 2 438 671 | – | – | – |
| 1929 | 2 533 229 | – | – | – |
| 1930 | 98,3% | 0,5% | 1,2% | 2 547 294 |

While no demographic data before the census in 1940 are completely reliable or accurate, especially concerning the numbers of Africans, the data provide valuable information concerning Angolan demographics in the forty-year period surveyed, i.e. the ratio was, in the first three decades of the 20th century, unfavourable to *mestiços*. Black Africans continued to constitute the overwhelming majority of the population.

Due to the increased numbers of white metropolitan settlers in Angola and the growing association between individuals' knowledge of Portuguese and their degree of "civilization", it is likely that the Afro-Portuguese had to improve their mastery of the language, but in the early

1920s Kimbundu was still the most widely used language in Luanda and the surrounding interior areas:

> ... a utilização da língua portuguesa pela minoria social de ascendência luso-africana não impediu que essa elite angolense praticasse o bilinguismo nas suas relações sociais. Porém, em ambiente familiar africano, falar português não era prática generalizada nem sequer em Luanda, apesar de, um século atrás, Sousa Coutinho ter proibido aos *moradores*, criarem seus filhos na "língua ambunda" como era corrente no século XVII (Freudenthal 2001:415-6).

The maintenance of Kimbundu as the mother tongue of the Afro-Portuguese elite at a time when there was a significant increase in number of white settlers in Angola seems to support the tentative hypothesis formulated above that structural borrowing in Kimbundu is likely to have intensified in the early 20th century.

However, while an increased number of white settlers in Angola was common to the colony as a whole, the demographic data presented in table 3 obscures the fact that this increase had specific demographic dynamics in particular regions. Hence, considering that the influx of white settlers was mostly felt in the cities, table 4 presents Amaral's (1962) estimate concerning the demographic population ratios in the main Portuguese urban centers in Angola in 1930.

**Table 4: The population in the cities of Angola according to race in 1930 (adapted from Amaral 1962)**

| CITIES | WHITES | *MESTIÇOS* | BLACKS | OTHER | TOTAL |
|---|---|---|---|---|---|
| Luanda | 12% | 11% | 77% | 0,04% | 50 588 |
| Malanje | 13,7% | 11,5% | 74% | 0,32% | 4 340 |
| Benguela | 8,8% | 8,4% | 82,7% | 0,03% | 12 674 |
| Lobito | 10% | 4,2% | 85,8% | – | 11 970 |
| Nova Lisboa | 17% | 8% | 74,4% | – | 14 064 |
| Silva Porto | 10,2% | 8,3% | 82% | – | 4 218 |
| Sá da Bandeira | 39,3% | 6% | 54,9% | – | 7 692 |
| Moçâmedes | 42,7% | 8,8% | 48,5% | – | 4 558 |
| TOTAL | 15,2% | 9,1% | 75,7% | – | 110 104 |

While the data on table 4 are only estimates, they indicate two clear demographic patterns. First, whereas in the coastal cities of Luanda, Benguela, Lobito and Nova Lisboa and in the interior cities of Malanje and Silva Porto the ratio of whites to mestiços is balanced, in the southern cities of Sá da Bandeira and Moçâmedes the former clearly outnumber the latter. Second, Sá da

Bandeira and Moçâmedes were also the only cities where the ratio of whites to blacks was relatively balanced.

The main reason for the reduced number of mestiços in Moçâmedes and Sá da Bandeira derives from the different types of exploration and settlement that were enforced in that region. In fact, whereas Luanda, Benguela, Lobito, Nova Lisboa, Malanje and Silva Porto were always dominated by trade, Sá da Bandeira and Moçâmedes were the scene of the first serious efforts to settle a significant number of whites. Consequently, the type of settlement that prevailed in the north was that of the trading post and the *presídio*, whereas in the south it was that of agricultural settlements (Amaral 1962:31-2). Moreover, the fact that the settlers arriving in the south came in relatively large numbers, usually arriving at the same time and from the same geographical areas (i.e. Pernambuco, Algarve and Madeira) contributed to the development of a stronger sense of community, which decreased the likelihood of miscegenation (*ibid.* 34).

In the light of this setting it is likely that the more balanced proportions of whites vs. Africans (both mixed-race and blacks) in Moçâmedes and Sá da Bandeira also corresponded to a more balanced ratio of native versus non-native speakers of Portuguese, both in the cities themselves and in the surrounding agricultural and penal settlements. However, at least at the start of the 20th century, this is not likely to have resulted in a more generalized use of Portuguese among the African populations. The main reason for this was the fact that, as a consequence of the type of settlement, racial segregation was much more deeply rooted in Moçâmedes and Sá da Bandeira than in any other place in Angola. If to this we add the fact that, unlike Luanda and the cities in the north, the black population in Moçâmedes and Sá da Bandeira was linguistically very heterogeneous, one comes up with a sociolinguistic setting in which a very clear-cut division is established between the area of influence of Portuguese and that of the Bantu languages.

In short, despite a significant increase in the number of white Portuguese settlers in Angola, which had different dynamics in the north and south of the country, one can state that the early 20th century in the colony was characterized by the more widespread use of the Bantu languages as opposed to Portuguese.

### 3.6.2. From the 1950s to 1975

The sociolinguistic setting described in the previous section was to remain unaltered until the 1950's, when the arrival of a growing number of white Portuguese settlers significantly altered

the demographics of the colony. The racial composition of the population in Angola is given in table 5 below:

**Table 5: Racial composition of the Angolan population (1940-1970) (based on Bender 2004:71)**

| | 1940 | 1950 | 1960 | 1970 |
|---|---|---|---|---|
| **Africans** | 98,1% | 97,4% | 95,3% | – |
| **Whites** | 1,2% | 1,9% | 3,6% | 5,1% |
| ***Mestiços*** | 0,75% | 0,72% | 1,1% | – |

The data above confirms the demographic trends that were already noticeable in the early 20th century in that it shows that Africans never constituted less than 95% of the population and that the proportion of *mestiços* in the Angolan society remained stable in comparison to the tremendous influx of white settlers. However, despite the fact that the white population of Angola nearly doubled every ten years, in the overall picture of the colony it never represented more than 5,1% of the total population.

The increase in the arrival of Portuguese settlers in the 1940s and 1950s was mostly due to the enforcement of the old policy of establishing big agricultural settlements in the interior, i.e. *povoamento planeado.* In the 1940s this policy had not produced any significant results, essentially due to the fact that most settlers preferred to work in commerce in the cities than in agriculture in the interior. In the 1950s the *povoamento planeado* produced better results, especially because, for the first time, the metropolitan government appeared to have the financial means to support these agricultural settlements (Bender 2004:185). The most significant of these in the 1950s were Cela (in the central plateau) and Matala (along the River Cunene). The social setting in these rural colonies is not likely to have favoured much contact between the Portuguese and the Africans as the latter were not only removed from their lands but also forbidden from working in these settlements (*ibid.* 188).

However, in the 1960s, with the beginning of the first pro-independence conflicts in Angola, the metropolitan government was forced to enforce a policy of settlement that included more black settlers. The first step for this new policy was the abolition of the *Estatuto do Indígena* in 1961 and the sending of some Cape Verdean free workers to the new multiracial agricultural settlements in the interior (Bender 2004:192). However, this was insufficient to secure the success of these settlements, as most settlers, both European and Cape Verdean, while continuing in Angola, left the rural areas for the urban centers on the coast. Hence, it is only

natural that, as shown in tables 6 to 8[23], the racial composition of the population in the main cities also underwent significant changes:

**Table 6: The population in the cities of Angola (1940) (adapted from Amaral 1962)**

| CITIES | 1940 | | | | |
|---|---|---|---|---|---|
| | WHITES | *MESTIÇOS* | BLACKS | OTHER | TOTAL |
| Luanda | 14,7% | 10,1% | 75,2% | 0,04% | 61 028 |
| Malanje | 16,3% | 15,5% | 67,8% | 0,3% | 5 299 |
| Benguela | 9,6% | 7,3% | 83% | 0,03% | 15 243 |
| Lobito | 11,9% | 4% | 84% | 0,007% | 13 592 |
| Nova Lisboa | 19,7% | 8,9% | 71,4% | 0,006% | 16 288 |
| Silva Porto | 14,2% | 8,3% | 77,5% | – | 4 671 |
| Sá da Bandeira | 39,4% | 5,5% | 55% | – | 8 521 |
| Moçâmedes | 44% | 8,8% | 47,2% | – | 4 926 |
| TOTAL | 17,2% | 8,8% | 74% | – | 129 568 |

**Table 7: The population in the cities of Angola (1950) (adapted from Amaral 1962)**

| CITIES | 1950 | | | | |
|---|---|---|---|---|---|
| | WHITES | *MESTIÇOS* | BLACKS | OTHER | TOTAL |
| Luanda | 14,6% | 6,9% | 78,4% | 0,04% | 141 647 |
| Malanje | 16,8% | 7,3% | 75,9% | – | 9 473 |
| Benguela | 22,8% | 5,7% | 71,5% | – | 14 690 |
| Lobito | 17% | 0,2% | 81% | – | 23 897 |
| Nova Lisboa | 16,8% | 4,2% | 79% | 0,03% | 28 296 |
| Silva Porto | 14% | 4,2% | 82% | 0,01% | 8 840 |
| Sá da Bandeira | 53,2% | 4,7% | 42,1% | – | 11 654 |
| Moçâmedes | 41,3% | 4,4% | 54,3% | – | 8 576 |
| TOTAL | 18,4 | 5,6% | 75,8% | - | 247 073 |

[23] The data referring to the years 1940 and 1950 are based on the Census that were carried out in those years. The data referring to 1955 are an estimate.

**Table 8: The population in the cities of Angola (1955) (adapted from Amaral 1962)**

| CITIES | 1955 | | | | |
|---|---|---|---|---|---|
| | WHITES | *MESTIÇOS* | BLACKS | OTHER | TOTAL |
| Luanda | 18% | 6% | 75,8% | 0,05% | 189 590 |
| Malanje | 16,6% | 5,3% | 78,2% | – | 12 815 |
| Benguela | 31,7% | 5% | 63,3% | – | 15 399 |
| Lobito | 20,2% | 1,3% | 78,5% | – | 31 630 |
| Nova Lisboa | 15,4% | 3% | 81,6% | 0,03% | 37 381 |
| Silva Porto | 13,7% | 3% | 83,4% | 0,008% | 12 146 |
| Sá da Bandeira | 59,6% | 4,2% | 36,3% | – | 13 867 |
| Moçâmedes | 60,2% | 4,9% | 35% | – | 7 185 |
| TOTAL | 21,1% | 5% | 73,9% | – | 320013 |

If one compares the total number of whites in the cities in 1940 (i.e. 22,289), 1950 (i.e. 45,453) and 1955 (i.e. 67,642) with the total numbers of whites in the colony as a whole in those years, respectively, 44,083; 78,826 and 109,568 (Amaral 1962), it is clear that the majority of the Portuguese preferred to settle in the urban centers, especially those along the coast (cf. the reduced number of white settlers in Malanje and Silva Porto in comparison to the other cities). Here, they occupied the positions previously held by the Afro-Portuguese elite in the public sector as well as in the economy and politics, and they enforced a climate of discrimination and repression against black and Afro-Portuguese Angolans that made the ability to speak Portuguese *a sine qua non* for any possibility of social advancement or even escape from the system of forced labor. The *Estatuto do Indígena* had been revoked but assimilation was still a key factor in Portugal's colonial policy in Africa.

The majority of Portuguese settlers were still mostly undereducated and poor; hence, most of them actually joined the large masses of Africans in the poor neighborhoods of Luanda (i.e. *musseques*) and other coastal cities, where they took the unskilled jobs available. Their low social status, however, did not correspond to a more humane attitude towards the Africans. The Colonial propaganda during the early 1900s and throughout the *Estado Novo* was far too deeply rooted in racism to allow it.

In fact, the sociolinguistic setting in the cities was more complex than in previous centuries. In fact, after 1961 the large numbers of rural Africans who did not cross the border into Zaire to

escape the conflict flocked to Angolan cities along the coast in search of better living conditions. This led to the concentration of large numbers of Africans from different linguistic backgrounds in the *musseques*. Due to their contact with whites, they acquired some degree of proficiency in Portuguese, but they still made up the majority of the population (according to Kasack 1996:72, the Portuguese constituted only 3,7% of the total population of the *musseques*). Hence, Africans were able to continue using their mother tongues. In the light of this sociolinguistic setting, primary dialect leveling is likely to have occurred as a consequence of different regional and social varieties of both Kimbundu and Portuguese coming into contact in the *musseques*.

In the interior, the population remained almost immune to Portuguese influence until the 1970s. However, in response to the growing influence of the nationalist movements in Angola, the Portuguese government invested heavily in increasing its presence in the interior and throughout the early 1970s the Portuguese military grouped much of the population in the interior, especially in the east, in the *aldeamentos*, i.e. "vast villages often surrounded by barbed wire, where previously dispersed Africans populations were kept together" (Bender 2004:264-5). One of the consequences of this policy was the mass exodus of Africans to neighboring countries, especially Zaire, which in turn led to the need to import labor from other provinces. Hence, in the *aldeamentos*, as in the urban centers, the sociolinguistic setting included the presence of large numbers of Africans from different linguistic backgrounds.

However, Portuguese is not likely to have been chosen as a lingua franca in these settings. Although they eventually led to the acquisition of some degree of proficiency in Portuguese, this does not seem to have been significant, since as late as the eve of Angola's independence from Portugal in 1975 the majority of the population in the interior was still not proficient in Portuguese and only a minority actually used it frequently. This is shown in Tables 9 and 10 below:

**Table 9: Level of proficiency in Portuguese (rural areas in Angola) [Heimer 1974:75 (vol. 2), cited in Bender 2004:353]**

| | |
|---|---|
| Perfect | 0,1% |
| Relatively reasonable | 0,4% |
| Reasonable | 16% |
| Rudimentary | 24% |
| No proficiency | 59% |

**Table 10: Frequency of use of Portuguese (rural areas in Angola) (*ibidem*)**

| | |
|---|---|
| Habitual | 0,1% |
| Some frequency | 0,8% |
| Very rarely | 31% |
| Never | 59% |

## 3.7. Summary and Conclusions

The early contacts in the Congo were mainly between the Portuguese and the Congolese nobility, and the use of Portuguese is likely to have been limited to the official communication between them. Consequently, the general population is likely to have continued speaking their mother tongues. As a consequence of the limited number of white settlers in the colony, especially women, the Portuguese were only capable of securing key trading posts along the coast and the River Kuanza, mostly by means of their mixed-race children.

In the Ndongo the Portuguese adopted a more expansionist rather than diplomatic policy. However, due to the inherent conditions in Angola (e.g. diseases, wars, harsh tropical weather) and their inability to attract enough settlers, especially women, to the newly founded colony, the Portuguese had to marry local women and to rely either on their mixed-race children or the native traders to penetrate further into the interior. Hence, Kikongo and later Kimbundu were the languages most widely spoken by both the white and black population in 16th century Angola.

The Portuguese dependence on their Kimbundu-speaking mixed-race children or on Africans continued throughout the 17th and 18th centuries, especially as a consequence of the intensification of the slave trade with Brazil and the Americas. This sociolinguistic situation remained practically unchanged until the first half of the 20th century, largely due to such factors as the resistance of Angolans to the Portuguese presence, the inhospitable conditions of the Angolan interior, an ineffective Portuguese administration until the 19th century, the failure of several methods of mainland colonization (e.g. establishment of penal colonies in the interior during the 1880s and 90s and the planned colonization in the 20th century), the delay in the construction of adequate roads and railway lines, the preference of the Portuguese for Brazil as a destination for emigration and their preference for settling in Angola's urban centers rather than in rural areas (Bender 2004). In short, well into the 20th century the Portuguese language in Angola was restricted to the major urban centers of the coast and to the educated elite.

This analysis of the sociolinguistic setting in Angola from the early days of contact to its independence from Portugal is consistent with the hypothesis that a Portuguese-based pidgin or creole never became established in Angola. Hence, the structural features of contemporary AVP are most likely to be the result of a process of language shift that did not begin in earnest until the mid-twentieth century. In fact, it was not until then that the conditions were met for the population, especially the bilingual Afro-Portuguese in the cities, to have enough contact with Portuguese or enough motivation to learn it to start the process of language shift. Several factors would seem to account for this. First, the great influx of white settlers put a significant part of the Angolan population in contact with Portuguese for the first time. Second, it was only during this period that most Angolans had enough motivation to learn Portuguese, either because of the prospect of upward social mobility or because of their need to cope with coexisting in the same community with other populations speaking different indigenous mother tongues.

The years that followed Angola's independence from Portugal reinforced Angolans' need to shift to Portuguese, not only because it was chosen as the official language and consequently the language of the state, media, culture, religion and the military, but also because the civil war that followed it brought a growing number of refugees to the cities along the coast, especially to Luanda, where "power and status depended on a firm grasp of Portuguese" (Birmingham 2002:157).

# Chapter 4

## Contact-induced features in the AVP Noun Phrase

### 4.0. Introduction

The aim of this chapter is to determine how the morphology of the noun phrase (NP) in AVP indicates the degree of contact-induced restructuring that Portuguese underwent in Angola, i.e. the extent of the retention of its morphosyntax as opposed to the introduction of substrate and interlanguage features. Hence, building on the theoretical framework proposed in Holm (2004), this chapter analyzes four morphosyntactic factors that can distinguish partially restructured vernaculars (PRV[s]) "from the standard variety of their source language" (*ibid.* 140), i.e. (i) number and (ii) gender marking and agreement, (iii) possession, and (iv) the word order and case of person markers.

In order to evaluate the influence of the languages in contact with Portuguese in the restructuring of AVP's noun phrase, sections 4.1 and 4.2 describe the noun phrase of EP and of the relevant Bantu languages, respectively regarding their internal structure, number and gender marking and agreement, possession, and the word order and case of person markers. Building on the insights gained in these sections, section 4.3 provides a parallel description and analysis of AVP's noun phrase. Finally, section 4.4 evaluates how the noun phrase in AVP indicates the degree of contact-induced restructuring that Portuguese underwent in Angola.

All of the data on the Bantu languages and the bulk of that on EP and AVP are drawn from existing literature. Concerning AVP, this is not unproblematic as the available data is exclusively drawn from written texts, so it is not clear whether it reflects the actual speech of most Angolans. Therefore, these data will be supplemented with the oral data from the present author's fieldwork in the interior of Angola, which, while limited, will hopefully provide some initial insights into the similarities and differences between the depiction of AVP in the literature and its actual use by speakers in everyday life.

## 4.1. The noun phrase in European Portuguese

Sections 4.1.1 to 4.1.4 below provide a description of the structure of the EP noun phrase concerning the four morphosyntactic phenomena mentioned above. However, *European Portuguese* is used here as an umbrella term to refer to the different varieties (both regional and social) spoken in Portugal, which includes but also transcends *standard European Portuguese* (SEP). This perspective is different from that of most works on AVP surveyed in chapter 2 (cf. sections 2.1.1 to 2.1.5), since these seem to SEP as the only "correct" variety of Portuguese and so tend to analyze the structure of AVP against the structure of this variety alone (i.e. the variety spoken by the educated speakers in the Coimbra-Lisbon region – cf. Cunha & Cintra 1995:10). All the other European varieties seem to be taken as deviating from this norm.

In contrast, in this thesis, all the European regional and social varieties of Portuguese are treated in the same way and are hence grouped together under the term *European Portuguese*[1]. Several reasons justify the use of the term in this broader sense as well as the analysis of the surveyed morphosyntactic phenomena in varieties and registers of EP other than the written or oral educated standard. Firstly, the term *variety* is used here in the sense of Peres & Móia (1995:34):

> Uma *variante* (...) de uma língua distingue-se pela associação do núcleo de características centrais dessa língua – lexicais, sintácticas, fonológicas – a um conjunto de características particulares envolvendo um ou mais destes níveis. Naturalmente, estas características têm de apresentar alguma estabilidade ao longo de um período razoável de tempo e, acima de tudo, têm de ser sustentadas por uma comunidade linguística minimamente representativa. (*my emphasis*)

Hence, features that are traditionally defined as deviations from the standard are defined here as part of the norm of specific varieties of EP, which means that a particular utterance may be incorrect according to the norm of the standard but correct according to the norm of a given variety. The register used (e.g. formal, informal, written, oral, etc) is also relevant in determining the adequateness of a given utterance. Only the features that conform to Peres & Móia's definition below are defined as true deviations from the norm (i.e. "errors" in traditional terminology) and are therefore preceded by an asterisk (*):

---

[1] The study of EP varieties has mostly focused on the description of regional varieties (cf. Vasconcellos 1901; Boléo & Silva 1961). The most recent dialectological classification of EP is that of Cintra (1971, 1983), which divides the regional varieties of EP into four groups: *dialectos galegos, dialectos setentrionais, dialectos centro-meridionais, dialectos das ilhas atlânticas*. The detailed description of the features of each of these groups of dialects transcends the scope of this thesis, so the reader is advised to consult the original works of the authors mentioned above. Nonetheless, should any feature of a specific regional variety of EP be relevant to the understanding of the restructuring of AVP, this will be indicated in the text.

> ... as construções, ou realizações fonéticas – que para nós configuram um desvio linguístico têm de obedecer a pelo menos duas condições: (i) constituírem rupturas com o subsistema ou variante de que é suposto fazerem parte; e (ii) não serem integradas – pelo menos, plenamente – pela comunidade linguística de suporte. (Peres & Móia 1995:41)

Secondly, the available data on the socio-economic background of the great majority of Portuguese settlers in Angola (i.e. mainly poor and uneducated) indicate that SEP could hardly have been the target variety for the L2 learners there (cf. chapter 3). Hence, I agree with Perl's assertion that "les facteurs linguistiques des variantes nationales en évolution d'un niveau quelconque du portugais en Afrique ne peuvent être comparés qu'à des éléments corrélatifs de même niveau du portugais européen ou brésilien" (1989b:14). Finally, comparing the structure of AVP to varieties of EP other than the standard also enables a clearer evaluation of the origin of the features of AVP's noun phrase surveyed here, i.e. whether these are the result of contact or of the natural drift of Portuguese[2], by reducing the probability that features of everyday informal speech of EP speakers be erroneously considered typical of AVP simply because they are ungrammatical or awkward in SEP.

### 4.1.1. Internal structure and word order

Concerning the internal structure of EP noun phrases, Mira Mateus *et al* (2003:325-70) state that the head noun can occur alone as in (1) or it can be accompanied by determiners (i.e. articles, demonstratives and possessives), as in (2) to (4), quantifiers (i.e. indefinites and numerals), as in (5) to (6), and/or modifiers (i.e. adjectives, prepositional phrases, relative clauses), as in (7) to (9):

| | | |
|---|---|---|
| (1) | **Angola** é fantástica! | 'Angola is fantastic!' |
| (2) | **O** debate foi interessante. | 'The debate was interesting.' |
| (3) | **Esse** debate foi interessante. | 'That debate was interesting.' |
| (4) | **O vosso** debate foi interessante. | 'Your debate was interesting' |
| (5) | **Alguns** debates foram interessantes. | 'Some debates were interesting.' |
| (6) | Houve **dois** debates. | 'There were two debates.' |
| (7) | Debates **interessantes** deveriam ter muito público. 'Interesting debates should have a good audience.' | |
| (8) | Raramente há debates interessantes **na televisão**. 'There are seldom interesting debates on television.' | |

[2] The discussion over the role played by internal as opposed to external factors in language change dates back to the late 19th century (cf. Chapter 2, section 2.2.2). Today, it is still one of the most disputed issues in the existing literature on the origin and synchronic structure of those language varieties which, like BVP, have fallen between "the cracks of theory, being neither unrestructured overseas varieties nor fully restructured creoles" (Holm 2004:144).

(9) O debate **que vi ontem** foi interessante.
'The debate I watched yesterday was interesting.'

The different elements in the NP conform to specific rules concerning the position of the non-nuclear elements in relation to the head noun (A) and the position of non-nuclear elements in relation to other non-nuclear elements in the NP (B). Hence:

**(A)**

**a.** *Definite and indefinite articles always precede the head noun:*

(10) Ela viu **a** carta. 'She saw the letter.'

(11) Ela viu **uma** carta. 'She saw a letter.'

**b.** *Demonstratives typically precede the head noun but they can also occur after the head noun in exclamations:*

(12) Que vida esta!

**c.** *Possessives typically precede the head noun (13), but they can follow the head in some exclamations or when preceded by a quantifier (14-15):*

(13) O **meu** trabalho de campo em Angola foi muito produtivo.
'My fieldwork in Angola was very productive.'

(14) Filhos **nossos** não farão isso! 'Our children will not do that.'

(15) Vi **sete** amigos **meus**. 'I saw seven friends of mine.'

**d.** *Adjectives can follow or precede the head noun depending on the pragmatic intention of the speaker (i.e. restrictive or subjective):*

(16) Um homem **pobre**. i.e. A man who is poor.

(17) Um **pobre** homem. i.e. An unfortunate man.

**(B)**

**e.** *Definite and indefinite articles never co-occur with demonstratives:*

(18) * Ela viu **a esta** carta. * 'She saw the this letter.'

(19) * Ela viu **uma esta** carta. * 'She saw an this letter.'

(20) Ela viu **esta** carta. 'She saw this letter.'

**f.** *Indefinite articles can never follow quantifiers:*

(21) * **Duas umas** cartas. * 'Two an letters.'

(22) * **Todas umas** cartas. * 'All an letter.'

(23) **Duas** cartas. 'Two letters.'

(24) **Todas** as cartas. 'All the letters.'

g. *Demonstratives can either occur as determiners or on their own as pronouns:*

(25) Este professor chegou atrasado. 'This teacher is late.'

(26) Este chegou atrasado. 'This one is late.'

h. *Possessives follow the definite articles, demonstratives and some quantifiers:*

(27) **Esse teu** comportamento é irritante. 'This behaviour of yours is annoying.'

(28) **Os nossos** amigos franceses chegam amanhã. 'Our French friends arrive tomorrow.'

(29) **Todos os nossos** amigos. 'All our friends.'

### 4.1.2. Number and gender marking and agreement

Agreement is used here to refer to a "systematic covariance between a semantic or formal property of one element and a formal property of another" (Steele 1978:610 quoted in Corbett 1994:55)[3]. Hence, it implies the existence of a controller (i.e. the element that determines the agreement, typically nouns or pronouns) and a target (i.e. the elements whose form is determined by agreement, typically adjectives, demonstratives, articles, possessives, numerals, relative pronouns, adverbs, adpositions, complementizers). The main functions of agreement are to introduce redundancy, so that it is easier to reconstitute the original message should part of it be lost, and to keep track of referents in a discourse (Corbett 1994:55-7).

In the noun phrase in EP only two agreement categories are morphologically encoded, i.e. number and gender[4]. However, the nature of this encoding is different for each category. As far as number is concerned, EP distinguishes between two degrees of numerosity: singular (one item) and plural (more than one item)[5], the last of which is the only degree which is formally marked, by adding the plural suffix marker *–s* to the stem of the noun. Number agreement is marked by adding this suffix to all the variable targets in the NP (i.e. determiners, quantifiers[6] and adjectives). Hence, number marking and agreement in EP are inflectional mechanisms, as they are both obligatory (i.e. they affect all variable nouns and non-nuclear elements in the NP)

[3] Many other definitions of *agreement* can be found in the extensive literature on the topic, but there is no generally accepted definition of the term for different linguists use different definitions according to their different theoretical backgrounds. Hence, I have chosen to use the one that is most often quoted in the literature. For a comprehensive view of the existing literature on *agreement* cf. the *Surrey Morphology Group Agreement Web Site* at http://www.surrey.ac.uk/LIS/SMG/projects/agreement/.

[4] Other *agreement categories* in the world's languages include *case*, *person* and *definiteness*. The latter does not exist in EP and the two former are lexically rather than morphologically marked.

[5] For a brief overview of number systems that allow for other degrees of numerosity (i.e. dual, trial, paucal, multiple, etc) see Cruse (1994). For a detailed cross linguistic analysis of the different typologies of number systems see Corbett (2000).

[6] The only numerals that are marked for number in EP are cardinal numerals ending in <ão> ("Um milhão" vs. "Dois milhões"), ordinal numerals ("O primeiro" vs. "Os primeiros") and fraction numerals, which agree with the cardinal numbers that indicate the number of parts ("**Um** terço" vs. "**Dois** terços") (Cunha & Cintra 1995:368-371).

and systematic (i.e. they affect all variable nouns and non-nuclear elements in the NP in the same way).

(30) Aquele gato branco 'That white cat.'

(31) Aqueles gatos brancos 'Those white cats.'

As far as gender is concerned, EP distinguishes between two values only, i.e. feminine and masculine, but unlike number, gender is not an inflectional category in EP. On the one hand, gender contrasts are not obligatory, i.e. only variable adjectives (cf. "novo/nova"; "inovador/inovadora") and nouns referring to animate beings (cf. *menino/menina* 'boy/girl'; *gato/gata* 'cat') contrast in gender. On the other hand, gender marking is not systematic, as different mechanisms are used to mark it, i.e. lexical mechanisms (cf. *novo/nova* 'new'; *menino/menina* 'boy/girl; *homem/mulher* 'man/woman') or morphological mechanisms such as derivation (cf. *conde/condessa* 'count/countess') and composition (*águia-macho/águia-fêmea* 'male eagle/female eagle') (Villalva 2000:219).

Hence, Villalva (2000:212-45) refuses the traditional definition of the suffixes *–a* and *–o* as inflectional markers of feminine and masculine, respectively[7]. She argues that these suffixes are morphological specifiers of the stem, i.e. *índice temático*, that enable the grouping of both nouns and adjectives into thematic classes similar to the Latin declensions. This is a highly complex system because nouns and adjectives are included into a particular thematic class depending on three factors: (a) the phonetic nature of the specifier, which establishes a distinction between *theme –a, theme –o, theme –e, theme Ø* and athematic words; (b) the existence of one as opposed to two forms for each gender, which distinguishes between variable and invariable nouns and adjectives; and (c) the gender of each word, which distinguishes between feminine and masculine nouns and adjectives.

Villalva (2000) identified 15 nominal thematic classes and 6 adjectival thematic classes, but in her contribution to Mira Mateus *et al* (2003:917-38) regarding the basic morphological structure of words in EP, the number of classes for nouns and adjectives is extended to twenty-three and nine, respectively, as shown in Tables 11 and 12:

[7] This view seems to be shared by Bechara (2002:132), who states that "a inclusão num ou outro género depende directa e essencialmente da classe léxica dos substantivos".

**Table 11: Nominal thematic classes in European Portuguese (adapted from Mira Mateus *et al* 2003:924)**

| THEMATIC CLASSES (variable nouns) | | | | | |
|---|---|---|---|---|---|
| | *-a* | *-o* | *-e* | *- Ø* | *athematic* |
| MASC. | poeta | aluno | infante | apresentador | avô |
| FEM. | aluna | -- | -- | imperatriz | avó |
| **THEMATIC CLASSES (invariable nouns)** | | | | | |
| MASC. | patriarca<br>mapa | ídolo<br>livro | abutre<br>dente | furriel<br>mar | tatu<br>café |
| FEM. | mosca<br>casa | tribo | gente<br>semente | variz | manhã |
| MASC. / FEM. | artista | modelo | agente | mártir | selvagem |

**Table 12: Adjectival thematic classes in European Portuguese (in Mira Mateus *et al* 2003:925)**

| THEMATIC CLASSES (variable adjectives) | | | | | |
|---|---|---|---|---|---|
| | *-a* | *-o* | *-e* | *- Ø* | *athematic* |
| MASC. | -- | claro | -- | falador | Bom |
| FEM. | clara | -- | -- | -- | Sã |
| **THEMATIC CLASSES (invariable adjectives)** | | | | | |
| MASC. / FEM. | careca | -- | leve | cortês | Ruim |

Hence, concerning gender marking in nouns, one can conclude that while nouns ending in *–a* tend to be feminine and those ending in *–o* tend to be masculine (Mira Mateus *et al* 2003:923), *-a* and *–o* cannot be interpreted as inflectional gender markers because, as Villalva (2000:243) points out, except for *theme –o* variable nouns, which are always masculine, it is possible to find (i) invariable *theme –o* nouns which are feminine (cf. *tribo* 'tribe') or both feminine and masculine (cf. *um/uma modelo* 'a model'); (ii) *theme –a* nouns which are masculine (cf. *poeta* 'poet'; *patriarca* 'patriarch'; *mapa* 'map') or both feminine and masculine (cf. *um/uma artista* 'an artist'); (iii) the gender of *theme Ø* and athematic nouns is completely arbitrary (cf. *o pente* 'comb' vs. *a lente* 'lens'; *o arroz* 'rice' vs. *a noz* 'nut').

Unlike the gender of nouns, gender is not inherent to adjectives and it has no referential content, i.e. gender in adjectives is exclusively grammatical because it is contextually determined by agreement with the noun. Nonetheless, two generalizations are possible: (a) all *theme –o* adjectives are masculine and (b) all *theme –e* adjectives are invariable (e.g. *leve* 'light'). Unlike verbs, the inclusion of nouns and adjectives into classes has no consequences for number inflection.

### 4.1.3. Possession

Possession is typically indicated by possessive determiners when referring to first and second person possessors, both singular [cf. (32) and (33)] and plural [cf. (34) and (35)]:

(32) A **minha** caneta. 'My pen.'
(33) A **tua** caneta. 'Your pen.'
(34) A **nossa** caneta. 'Our pen.'
(35) A **vossa** caneta. 'Your pen.'

The possessive determiners agree in both number and gender with the possessed, but not with the possessor:

(36) João, as **tuas** canetas? 'John, where are your pens?'
(37) Maria, os **teus** lápis? 'Mary, where are your pencils?'
(38) Meninos, a vossa mãe? 'Boys, where is your mother?'
(39) Meninas, o vosso pai? 'Girls, where is your father?'

With regard to word order, possessive determiners in EP typically precede the possessed noun. As illustrated in Mira Mateus *et al.* (2003:349-51), only in very specific contexts do the former follow the latter, namely when (a) the possessed noun is not preceded by an article or when it is used with an indefinite (e.g. *tens aí livros meus* 'you have books of mine there' or *tens aí alguns livros meus* 'you some books of mine there'); (b) when the possessed noun is preceded by an interrogative quantifier (e.g. *tens aí quantos livros meus?* 'how many books of mine do you have there?'); (c) in some exclamations (e.g. *filhos nossos não farão isso!* 'No child of ours will ever do that!); (d) when the head noun is preceded by a cardinal numeral (e.g. *vi sete amigos meus* 'I've seen seven friends of ours');

The agreement of the possessive determiners with the possessed but not with the possessor leads to ambiguity when the latter refers to the third person [cf. (40) and (41)]. In order to clarify this ambiguity the possessive determiners *seu(s), sua(s)* are often replaced with the amalgam of the

preposition *de* and the third person marker *ele* (M) or *elas* (F), in which case the preposition can never be omitted [cf. (42)]:

(40) José, Pedro levou o **seu** chapéu. (Bechara 2002:181)
'Joseph, Peter took his hat.' or 'Joseph, Peter took your hat.'

(41) José, Pedro levou o chapéu **dele**.
'Joseph, Peter took his hat.'

(42) * José, Pedro levou o chapéu _ele.

As stated in section 4.1.1, possessive determiners typically precede the head noun, except in a very limited set of contexts (cf. Cunha & Cintra 1995:320-1).

### 4.1.4. Case and word order of person markers

EP makes the classical distinction between "the speaker of an utterance, the addressee of that utterance and the party talked about that is neither the speaker nor the addressee" (Siewierska 2004:1), i.e. first, second and third persons, respectively. As far as the morpho-phonological form of person markers[8] is concerned, EP uses both independent or free forms and clitic forms. The former constitute "separate words and may take primary stress" (*ibid.* 16), whereas the latter form "a phonological unit with a word (their host) preceding it (enclitics) or following it (proclitics)" and are "written as separate words and are able to take, under some conditions, lexical stress" (*ibid.* 26)[9].

Concerning the function of person markers, EP distinguishes between subject and object forms, i.e. EP person markers contrast in *case*. Subject forms are exclusively free (i.e. *eu, tu, ele, ela, nós, vós, eles, elas*) whereas object forms are both free (i.e. *mim, comigo; ti, contigo; ele, ela, si; nós, connosco; vós, convosco; eles, elas*) and clitic (i.e. *me, te, o, a, lhe, nos, vos, os, as, lhes*).

As with most of the world's languages, person markers in EP do not mark person only. In fact all person markers in EP contrast in number, hence establishing an opposition between first, second and third person singular and first, second and third person plural; and third person subject and object pronouns, both singular and plural, contrast also in gender (i.e. feminine *vs.* masculine)[10]. In fact, cross-linguistically gender oppositions seem to be characteristic of third

[8] I follow Siewierska's (2004) option for the term *person marker* and *person form* in preference to *personal pronoun* as I agree that, indeed, "the term pronoun is open to a number of interpretations and even under the most liberal of these, not all grammatical markers of the category person are uncontroversially pronominal" (*ibid.* 2), e.g. EP expresses person not only by means of pronouns but also by means of nominal expressions (i.e. the so-called *courtesy pronouns*).

[9] For a full account of the typology of person markers concerning their morpho-phonological form, see Siewierska (2004:16-40).

[10] In addition to the grammatical categories of number, gender and case, person markers in EP also encode

(especially 3PL) rather than first or second person. According to Siewierska (2004:105) this is because:

> The genders of the first and second persons are typically self-evident to both the speech act participants and thus gender marking of the first and second persons is communicatively redundant. The gender of third persons, on the other hand, is not obvious. Though third parties may be present in the extra-linguistic context of the utterance, they need not be and in fact typically are not. Gender marking of the third person therefore helps the interlocutors to keep track of which third-person referent is being talked about.

It must be noted that first and second person clitic markers *me, te, nos, vos* can be used both as direct or indirect objects, but the third person is expressed by different markers depending on whether it is used as direct (i.e. *o, os, a, as*) or indirect object (i.e. *lhe, lhes*) (Cunha & Cintra 1995:302):

(43) Eu respeito-**o** muito. 'I have great respect for him.'

(44) * Eu respeito-**lhe** muito.

The free and clitic object forms are also in complementary distribution, as in (45), because the former typically either occur only after verbs that are followed by prepositions, as in (46) or in a very reduced set of syntactic contexts[11], as in (47) and (48):

(45) Ele viu-**a**. 'He saw **her**.'
Cf. * Ele viu **ela**. * 'He saw **she**.'

(46) Ele viu-**a**. 'He saw her' *vs.* Ele gosta **dela**. 'He likes her'

(47) *A ele* cumpria encher as guias. (Bechara 2002:174)

(48) Remeti os livros *a ele* e ao tio. (Bechara 2002:174)

Nonetheless, native speakers of SEP commonly use *ele, ela* instead of *o, a* as direct objects in their casual speech:

(49) Olha **ele**! 'Just look at him!' (Bechara 2002:175)

This use of *ele, ela* as direct objects is only possible in contemporary SEP for emphatic purposes[12], but its unmarked pragmatic use is attested in archaic Portuguese:

---

information concerning the social status of the speaker in relation to the speaker by means of the so-called courtesy pronouns (i.e. *pronomes de tratamento*). These will not be analyzed here as they imply agreement with the verb, which goes beyond the scope of this thesis.

[11] For a detailed list of these contexts cf. Bechara (2002:174).

[12] Person emphasis can also be expressed in EP by means of (a) the words *mesmo* and *próprio* (e.g. **Tu mesmo** disseste isso. 'You said that yourself'; **Eu própria** falarei com o João 'I will talk to João myself.' ) or (b) the invariable expression *que* (e.g. **Eu é que** vou falar com o João 'I am the one who will talk to João'; **Tu é que** vais falar com o João 'You are the one who will talk to João'; **Ele é que** vai falar com o João 'He is the one who will talk to João.'). Both strategies can only be used with subject person markers (Cunha & Cintra

> Na literatura antiga encontra-se por vezes o pronome *el* ou *ele* empregado também como acusativo ou complemento directo; assim, no *Livro de Esopo* (edição de Leite de Vasconcelos) lê-se a pág. 33 *que enforcariam ell.* Veja-se exemplo idêntico em documento do século XIII, publicado por P. de Azevedo na *Revista Lusitana*, VIII, pág. 39. (Nunes 1989:238)

Hence, the different paradigms of person markers in SEP can be systematized as shown in Table 13 below:

**Table 13: Person markers in European Portuguese**

| | INDEPENDENT MARKERS | | DEPENDENT MARKERS | |
|---|---|---|---|---|
| | *Subject* | *Object* | *Direct object* | *Indirect Object* |
| 1SG | eu | mim, comigo | me | |
| 2 SG | tu | ti, contigo | te | |
| 3 SG | ele, ela | ele, ela, si | o, a | lhe |
| 1PL | nós | nós, connosco | nos | |
| 2 PL | vós | vós, convosco | vos | |
| 3 PL | eles, elas | eles, elas, si | os, as | lhes |

When the direct or indirect object coincides with the subject, it is expressed by means of clitic markers traditionally referred to as reflexive pronouns. Except for the 3SG/PL marker *se,* first and second person reflexive markers are homophonous with the corresponding non-reflexive clitic object markers above, i.e. *me* (1SG), *te* (2SG), *se* (3SG/PL), *nos* (1PL) and *vos* (2PL):

(50) Eu magoei-**me**. 'I hurt myself.'

(51) Tu magoaste-**te** ontem. 'You hurt yourself yesterday.'

(52) Ele magoou-**se** ontem. 'He hurt himself yesterday.'

(53) Nós magoámo-**nos**. 'We hurt ourselves.'

(54) Vós haveis-**vos** magoado. 'You have hurt yourselves.'[13]

(55) Eles magoaram-**se** ontem. 'They hurt themselves yesterday.'

When the verb refers to a subject that is constituted by two beings or things, the reflexive markers *nos, vos*[14], *se* are called reciprocal person markers (i.e. reciprocal pronouns in traditional terminology):

(56) Carlos e eu abraçamo-**nos**. 'Carlos and I hug each other'

(57) José e António não **se** cumprimentam. 'José and António do not talk to one another.'

---

1995:289).

[13] The free form *vós* and the corresponding clitic form *vos* are very little used in EP, even in the speech of highly educated speakers. It is systematically replaced with the second person plural expression *vocês* (e.g. Vocês magoaram-**se**. 'You have hurt yourselves.')

[14] Cf. footnote 13 on the use of the second person markers in EP.

As far as the word order is concerned, EP clitic person markers typically appear in enclitic position, i.e. after the verb (Mira Mateus *et al* 2003:852). This is made clear by several studies which show that Portuguese children (cf. 58-59) and Portuguese youngsters (cf. 60-61) tend to display the enclitic pattern:

(58) não chama-**se** nada (M., 20 months old in Mira Mateus *et al* 2003:850)

(59) porque é que foste-**me** interromper (R, 29 months, *ibid.*)

(60) porque não apercebeu-**se que** ... (12 years old, written text *ibid.* 851)

(61) Também sabe-**se** que existe uma certa altura da criança
(undergraduate student, *ibid.* 852)

However, in certain syntactic, semantic and prosodic contexts, clitic person markers must occur either between the past participle auxiliary and the main verb (A), before the verb (i.e. proclitic position) as in (B) or "inside" the verb form (i.e. mesoclitic position) as in (C).

**(A)**

(62) A prenda foi-**nos** oferecida pela avó.
'The gift was given to us by our grandmother.'

(63) O Mário tem-**na** ajudado muito.
'Mário has helped her a lot.'

**(B)**

**(a)** *In negative sentences or phrases*:

(64) O João **não/nunca me** telefonou. 'João did not/never call me.'

(65) **Ninguém/nada o** demoveu. 'Nobody/nothing made him change his mind.'

**(b)** *In interrogative and relative clauses*:

(66) **Quem te** disse isso? 'Who told you that?'

(67) **Onde os** encontraste? 'Where did you find them?'

**(c)** *In subordinate clauses:*

(68) Quando eles **os** viram foram cumprimentá-los.
'When they saw them they went to talk to them.'

(69) **Se** eles **os** encontrassem, era óptimo. 'It would be great if they found them.'

**(d)** *With preverbal numerals and indefinite determiners:*

(70) **Todos se** levantaram. 'They all got up.'

(71) **Alguém a** viu? 'Has anyone seen her?'

(C)

(e) *With future and conditional verb forms:*

(72) Dir-**lhe**-ei como fazer isto. 'I will tell you how to do this.'

(73) Dir-**lhe**-ia para ir ao médico. 'I would tell you to see a doctor.'

The use of clitic person markers in enclitic position is in clear opposition to the typical position of these markers in 15th century EP, i.e. proclitic position. The change seems to have taken place in the 17th century, which leads Mira Mateus *et al* (2003:850) to conclude that the "inversão brusca dos padrões proclítico e enclítico pode ser atribuída a uma mudança na gramática do português ocorrida em meados do século XVIII, com várias consequências."

## 4.2. The noun phrase in Kikongo, Kimbundu, Umbundu and Cokwe

Sections 4.2.1 to 4.2.5 below describe the structure of the NP in Kikongo (KK), Kimbundu (K) and Umbundu (U) (i.e. the African languages that were in closer contact with EP along the Angolan coast) and Cokwe (CC) (i.e. the native language of the speakers interviewed during my fieldwork) with regard to the four morphosyntactic phenomena mentioned above (i.e. number and gender marking and agreement, possession and the word order and case of person markers). However, it is impossible to analyze these phenomena in isolation from the NP morphology of the Bantu languages in general because while they display many language-specific features, these phenomena cannot be fully understood without prior knowledge of the general rules that govern them, especially those concerning the complex system of nominal classification and agreement.

In fact, the most prominent morphosyntactic feature of the Bantu languages, and the one that best attests to their agglutinating character, is their complex system of noun classes. These constitute a "closed obligatory grammatical system", so "in most cases there is a limited, countable number of noun classes" (Aikhenvald 2003:21). It is estimated that noun classes may total twenty-three, although not all of them may occur in individual languages (Maw 1994:305).

One of the salient features of the Bantu noun class system is the fact that each noun typically takes a noun class marker (i.e. prefix). Therefore, what we call nouns in the Bantu languages consist of a prefix (i.e. class marker) plus a stem (Maw 1994:305)[15]. Each noun class typically correlates to a distinctive set of prefixes denoting a singular-plural opposition, so when a noun shares the same set of prefixes it is traditionally said to belong to the same noun class:

[15] This leads some bantuists to use the term nominal group instead (cf. Maw 1994; Mutaka 2000).

| | | | | | | |
|---|---|---|---|---|---|---|
| (74) | KK: | *m-ntu*<br>C1.SG-STEM | 'person' | vs. | *a-ntu*<br>C1.PL-STEM | 'people' |
| (75) | K: | *mu-tu*<br>C1.SG-STEM | 'person' | vs. | *a-tu*<br>C1.PL-STEM | 'people' |
| (76) | U: | *omu-nu*<br>C1.SG-STEM | 'person' | vs. | *oma-nu*<br>C1.PL-STEM | 'people' |
| (77) | CC: | *mu-tfu*<br>C1.SG-STEM | 'person' | vs. | *a-tfu*<br>C1.PL-STEM | 'people' |

These prefixes that indicate the grammatical class and number of the noun are traditionally referred to as primary prefixes, as opposed to the secondary prefixes, which add "something more lexical like: small, huge, amiable, despicable, etc" (Mutaka 2000:154), e.g. locatives, augmentatives, diminutives, etc. The secondary prefixes can either: (a) replace the primary prefix or (b) be prefixed to it.

Nonetheless, the key definitional property of the Bantu noun class system is that non-nuclear elements in the NP must agree in class and number with the head noun by means of agreement markers (cf. Aikhenvald 2003:21; Corbett 1995:45), i.e. each noun class correlates to a specific set of agreement markers denoting a singular-plural opposition. As in EP, the non-nuclear elements in the Bantu NP include determiners (i.e. articles[16], demonstratives and possessives), quantifiers (i.e. indefinites and numerals) and/or modifiers (i.e. adjectives[17], prepositional phrases, relative clauses)[18]. The agreement prefixes can be broadly divided into two types, i.e. "one form of agreement which appertains to lexical items (nouns, adjectives, epithets, and numerals) – in other word items that belong to an infinitive set - and a different form which appertains to grammatical items (possessives, demonstratives) – which belong to closed systems" (Maw 1994:306).

In the work of early Bantuists each individual prefix was given a number and each was considered a noun class on its own (i.e. typically odd numbered classes were singular and even numbered classes were plural). Hence, in (74) *m-* and *a-* would be markers for class 1 and class

[16] As far as articles are concerned, KK, K and U follow the general Bantu pattern, i.e. they only display definite articles. CC displays neither definite nor indefinite articles. The form of the definite article in KK, K and U is the same for both singular/plural and feminine/masculine. The form of the article in K and U is *o*, but in Kikongo there are three different forms: *o, e, a*. In all three languages the article precedes the class prefix, despite the fact that in particular phonological contexts it can be omitted. In contemporary Bantuist literature these forms are often referred to as *augments* (e.g. Lojenga 1994:303).

[17] The set of available adjectives in the African languages in general and in the Bantu languages in particular is highly reduced and can be listed in half a page (for a list of the available adjectives in the Angolan Bantu languages surveyed here cf. Nascimento 1893:26; Chatelain 1888-89:64; Tavares 1915:27; Maia 1957:16). Moreover, in most African languages, the concepts traditionally expressed by adjectives in most Indo-European languages "tend to take the shape of nouns or verbs" (Dimmendal 2000:171).

[18] Concerning the word order of these elements in the NP, it is not possible to establish a basic pattern because this is essentially language-specific and allows for extensive variation (Maw 1994:306).

2, respectively[19]. However, this traditional approach has two problems. First, it blurs the link between singular and plural classes, which is why in modern bantuist literature the traditional numbering system is maintained but it is each set of singular-plural prefixes and not each prefix individually that is taken to constitute a noun class, e.g. in (74) both *m-* and *a-* are class 1/2 markers[20]. Second, it assumes that the decisive factor in class assignment is the morphological form of the noun, i.e. each noun is included in a particular class depending on the set of prefixes it takes (cf. Corbett 1995:45-6). If this were true, since all non-nuclear elements in the NP must agree in class and number with the head noun, one would expect that the class of agreement markers would always match that of the prefixes on the noun, i.e. alliterative concord (Aikhenvald 2003:35), as in (78) and (79) below:

(78) SWAHILI: *Kikapu kikubwa kimoja kilianguka* (Corbett 1995:43)
7-basket 7-large 7-one 7-fell
'one large basket fell'

(79) SWAHILI: *Vikapu vikibwa vitatu vilianguka* (*ibid* 44)
8-baskets 8-large 8-three 8-fell
'Three large baskets fell'

However, we often find quirky agreement in the Bantu languages, i.e. there is a mismatch between the noun class marked on the noun itself and the agreement class it is assigned to (Aikhenvald 2003:38). For example, Swahili's *kifaru* 'rhinoceros' takes class 7/8 prefixes (i.e. *ki-* and *vi-*), but the non-nuclear elements in the NP of which it is the head take class 1/2 agreement markers:

(80) SWAHILI: *Kifaru m-dogo a-likuwa hapa*
rhinoceros C1-small C1-was here
A small rhinoceros was here.

Hence, the morphological form of the noun (i.e. the set of prefixes it take) is not the only class assignment criterion, as class assignment in modern Bantu is based on both morphological and semantic principles. This position is best illustrated by Corbett (1995), who suggests that class assignment in Bantu is primarily governed by semantic rules, with morphological rules playing a complementary role. Hence, semantic rules assign nouns to a particular class according to the semantic properties of their referents (e.g. all animate nouns belong to class x/z) regardless of their morphological class. Only the nouns that are not assigned a class by semantic rules are

[19] For a reproduction of the reconstructed Proto-Bantu noun prefix system see Mutaka (2000:151).

[20] The authors of the KK, K, U, and CC grammars used here seem to use a yet different system of noun class numbering, i.e. they do not number each prefix individually but rather identify each set of singular/plural prefix with a number (Arab or Roman). Hence, *m-* and *a-* in (74) are both class 1 prefixes.

then included in the class that corresponds to their morphological form[21].

Hence, noun classes have much in common with the grammatical category of gender, as they are "classes of nouns reflected in the behavior of associated words" (Hockett cited in Corbett 1995:1). Therefore, noun classes are most often referred to in contemporary bantuist literature as genders, i.e. "sets of nouns which take the same agreements" (Corbett 1995:45). According to this line of thought, the traditional designations *noun class* and *agreement class* are replaced with *controller gender* (i.e. sets into which nouns are divided) and *target gender* (i.e. agreement forms found in the targets), respectively (Corbett 1995:45). However, it must be noted that the grammatical category of gender in the Bantu languages goes beyond the traditional male/female distinction in most Indo-European languages; it includes other oppositions such as "human, animal, plant, body parts, mass nouns and liquids, abstracts, and others which are less transparent" (Williamson & Blench 2000:13). In fact, most Bantu languages do not morphologically express the male/female distinction (cf. section 4.2.2 for a description of the mechanisms used to express this distinction in KK, K, U and CC).

The authors of the available grammars of KK, K, U and CC are agreed on the importance of target genders, which they refer to as "partículas concordantes" or "prefixos concordantes", but, despite their awareness of several instances of mismatches between the class marked on nouns and the agreement prefixes on the agreeing elements as a consequence of the semantic properties of the noun (cf. Maia 1957:107), most authors assume that it is the prefix on the noun that determines its inclusion in a given class (cf. Chatelain 1888-1889:1). Consequently, there is no general description of the class assignment rules used in these languages. This is mainly because these works were mostly published in the late 19th century and early 20th century, well before it was clear to linguists that class assignment is governed by semantic and morphological rules. Moreover, the available data on these languages is insufficient to allow for any tentative description of their class assignment systems. Hence, tables 14, 15, 16 and 17 provide the noun class system of KK, K, U and CC using the conventions originally presented by Tavares (1915), Chatelain (1888-9), Nascimento (1893) and Martins (1990), respectively.

[21] According to Aikhenvald (2003:21), "for nouns with animate reference, semantic assignment often overrides morphological or other principles" of class assignment.

**Table 14: Noun class system in Kikongo (based on Tavares 1915; Santos 1962:33)**

| Noun Class | Class Markers | | Agreement Markers | | Semantics |
|---|---|---|---|---|---|
| | Sg | Pl | Sg | Pl | |
| 1 | *mu-*<br>*m-*<br>*n-* | *a -* | *u-* | *a-* | –*human* and *animate* beings. |
| 2 | *mu-*<br>*m-*<br>*n-*<br>∅ | *mi –*<br>*mim-*<br>*min-* | *u-* | *mi-* | –nouns whose SG prefix is *mu-* and which denote *irrational* or *inanimate* beings. |
| 3 | *ki –*<br>∅ | *i -* | *ki-* | *i-* | –nouns whose SG prefix is *ki-*.<br>–nouns which used to take the SG prefix *ki-*;<br>–*objects, instruments, languages, place names* and *qualities*. |
| 4 | *di –*<br>∅ | *ma -* | *di-* | *ma-* | –nouns whose SG prefix is *di-*.<br>–*massive* nouns, which take no singular prefix |
| 5 | *ku -* | *ma -* | *ku-* | *ma* | –nouns whose SG prefix is *ku-*.<br>–*body parts*.<br>–nominal verb forms. |
| 6 | *lu -* | *tu -* | *lu-* | *tu-* | –nouns whose SG prefix is *lu-*.<br>–*collective nouns* denoting *individual entities*. |
| 7 | *m-*<br>*n -* | *zim -* | *i-* | *zi-* | –nouns whose beginning with *m-* or *n-* followed by consonant.<br>–*borrowings*.<br>–*collective* nouns proper.<br>–*deverbal nouns*. |
| 8 | ∅ | *ma -* | *u-* | *ma-* | –nouns whose SG prefix used to be *u-* |
| 9 | *u -* | *u-* | *u-* | *u-* | –abstract nouns.<br>–non-abstract nouns beginning with *u-*. |
| 10 | *va -* | *mu -* | *va-* | *mu-* | –constituted only by a single noun (cf. example). |

**Table 15: Noun class system in Kimbundu (based on Chatelain 1888-1889)**

| Noun Class | Class Markers | | Agreement Markers | | Semantics |
|---|---|---|---|---|---|
| | Sg | Pl | Sg | Pl | |
| 1 | *mu-* | *a-* | *u-* | *a-* | –*human* and *animate beings* whose SG prefix is *mu-*. |
| 2 | *mu-* | *mi-* | *u-* | *i-* | –*inanimate* referents whose SG prefix is *mu-*.<br>–*trees* and *plants*. |
| 3 | *ki-* | *i-* | *ki-* | *i-* | –nouns whose SG prefix is *ki-* (including *people* and *animals*).<br>–*augmentative* nouns. |
| 4 | *ri-* | *ma-* | *ri-* | *ma-* | –nouns whose SG prefix is *ri-*. |
| 5 | *u-* | *mau-* | *u-* | *ma-* | –*abstract* nouns. |
| 6 | *lu-* | *malu-* | *lu-* | *ma-* | –*collective* nouns. |
| 7 | *tu-* | *matu-* | *tu-* | *ma-* | –[rare class] |
| 8 | *ku-* | *maku-* | *ku-* | *ma-* | –*deverbal* nouns. |
| 9 | ∅ | *ji-* | *i-* | *ji-* | –*borrowings*. |
| 10 | *ka-* | *tu-* | *ka-* | *tu-* | –*diminutive* nouns.<br>–nouns deverbal nouns. |

**Table 16: Noun class system in Umbundu (adapted from Nascimento 1893)**

| Noun Class | Class Markers | | Agreement Markers | | Semantics |
|---|---|---|---|---|---|
| | Sg | Pl | Sg | Pl | |
| 1 | *omu-*<br>*u-* | *oma-*<br>*a-* | *u-* | *va-*<br>*a-* | – *animate* beings. |
| 2 | *omu-*<br>*u-* | *omi-*<br>*ovi-* | *u-* | *vi-* | – *inanimate* beings, *vegetables*, etc. |
| 3 | *e-*<br>*i-* | *a-*<br>*ova-* | *li-* | *va-*<br>*a-* | – *proportions*, *body parts*, *abstract* nouns. |
| 4 | *ochi-* | *ovi-* | *chi* | *vi-* | – *instruments, animals, causes, functions.*<br>– *augmentatives.* |
| 5 | *o-* | *olo-* | *i-* | *vi-* | – *animals.*<br>– *deverbal nouns.*<br>– *borrowings.* |
| 6 | *olu-* | *olo-* | *lu-* | *vi-* | – *extension.* |
| 7 | *oka-* | *otu-* | *ka-* | *tu-* | – *small things* and *diminutives.* |
| 8 | *ou-* | *∅*<br>*au-* | *u-* | *va-*<br>*a-* | – *abstract* nouns. |
| 9 | *oku-* | *ova-* | *ku-* | *va-*<br>*a-* | – *body parts* and *verbal expressions.* |

**Table 17: Noun class system in Cokwe (adapted from Martins 1990)**

| Noun Class | Class Markers | | Agreement Markers | | Semantics |
|---|---|---|---|---|---|
| | Sg | Pl | Sg | Pl | |
| 1 | *mu-* | *a-* | *u / mu* | *a* | – *humans* or *personified things* whose prefixes are *mu-* (SG) and *a-* (PL). |
| 2 | *mu-* | *mi-* | *u* | *i* | – nouns whose prefixes are *mu-* (SG) and *mi-* (PL).<br>– all *trees* and *plants.* |
| 3 | *tchi-* | *i-* | *tchi* | *i* | – nouns whose prefixes are *tchi-* (SG) and *i-* (PL).<br>– some *augmentatives.* |
| 4 | *ri- / li-* | *ma-* | *ri / li* | *a* | – all nouns whose prefixes are *ri-/li-* (SG) and *ma-* (PL). |
| 5 | *u-*<br>*∅* | *maw-*<br>*u-* | *u* | *u* | – all *abstract nouns* that have *different forms* for *singular* and *plural.* |
| 6 | *lu-* | *malu-* | *u* | *a* | – nouns whose prefixes are *lu-* (SG) and *malu-* (PL). |
| 7 | *tu-* | *matu-* | *i* | *a* | – nouns whose prefixes are *tu-* (SG) and *matu-* (PL). |
| 8 | *ku-* | *maku-* | *u / ka* | *lu / a* | – nouns whose prefixes are *ku-* (SG) and *maku-* (PL). |
| 9 | *i-*<br>*∅* | *ji-* | *i* | *ji* | – mostly *borrowings.* |
| 10 | *ka-* | *tu-* | *ka* | *tu / a* | – nouns whose prefixes are *ka-* (SG) and *tu-* (PL).<br>– almost all *diminutives.* |

Before proceeding with the analysis of number and gender marking and agreement, possession and case and word order of person markers in the Bantu languages surveyed here, a final remark is in order concerning the class (i.e. gender) assignment of borrowed words, as bantuists are not agreed on the nature of this process; some support the hypothesis that special rules of class assignment are used, i.e. rules that are different from those used with native words. According to Corbett (1995:75-82), semantic analogy (i.e. the loanword takes the gender of a noun of similar meaning already in the language), unmarked gender (i.e. loans are assigned to the gender with the largest number of words in it) and automatic assignment (i.e. the language has a gender that includes loanwords alone) have been suggested as the key special processes of gender assignment to borrowings. Nonetheless, the most widely accepted theory is that "loanwords are assigned to a gender according to meaning and form, depending on the assignment system of the borrowing language. This suggests that the gender of loanwords is determined exactly like that of other nouns" (Corbett 1995:74). For example, in Swahili, two rules are used to assign gender to loanwords:

> ... the majority of nouns denoting animates are assigned to genders 9/10 and 5/6. The reason for these choices is that morphological classes 9/10 and 5/6 are precisely the ones which include nouns with no prefix in the singular [...]
>
> ... the initial segment of the borrowed word matches a prefix and the noun may enter the appropriate morphological class. (Corbett 1995:72)

As far as Portuguese loans to Bantu languages in Angola are concerned, both their inclusion into classes that have no prefix in the plural or whose prefix matches the initial segment of the loan have been attested in the literature (cf. Oliveira 1990:89).

### 4.2.1. Internal structure and word order in KK, K, U and CC

As far as the internal structure of the NP is concerned, KK, K, U and CC follow the general pattern of the Bantu languages. Hence, the NP may be constituted by the head noun alone or it may comprise the head noun plus determiners (i.e. articles, demonstratives and possessives), as in (81) to (83), quantifiers (i.e. indefinites and numerals), as in (84) to (85), and/or modifiers (i.e. adjectives, prepositional phrases, relative clauses), as in (86) to (88). For a matter of convenience and considering the uniformity of the NP structure in the four languages surveyed here, this is illustrated with examples from Kimbundu only:

(81) K: *o* *mu-hatu* (Chatelain 1888-1889:149)
the C1.SG-woman
'the woman'

(82) K: *mu-tu* ***iuná*** (*ibid* 27)
C1.SG-person that
'that person'

(83) K: *a-bik'* ***ami*** (*ibid* 15)
C1.PL-slave 1SG.POSS
'my slaves'

(84) K: *ngiji* ***i-engui*** (*ibid* 147)
C9.SG.river C9.AGR-other
'other river'

(85) K: *ma-longa* ***ma-tanu*** (*ibid* 148)
C4.PL-plate C4.AGR-five
'four dishes'

(86) K: *ki-mbiji* ***ki-onene*** (*ibid* 148)
C3.SG-fish C3.AGR-big
'big fish'

(87) K: ***mu*** *njibela* (*ibid* 149)
in C9.SG.pocket
'in/into the pocket'

(88) K: *O tu-veia* *tu-iadi* (Chatelain 1894:33)
the C10.PL-old women C10.AGR-two

***tu*** *u-a-sange*
whom 2SG-PST-meet
'the two old women whom you met'

Examples (84) to (88) show that the head noun is typically the leftmost element in the NP. Concerning the word order of the non-nuclear elements in the NP in relation to the head noun, the examples above show that only the definite article and prepositions can precede it. All the remaining non-nuclear elements occur after the head of the NP (i.e. demonstratives, possessives, indefinites, numerals, adjectives and relative clauses). Concerning the rules governing the position of non-nuclear elements in relation to other non-nuclear elements in the NP, only Martins (1990) provides a systematized account. Hence, (89) to (91) are exclusively drawn from Cokwe, despite the fact that the word order of the non-nuclear elements is quite similar in KK, K and U.

**(a)** Possessive adjectives always precede cardinal numerals:

(89) CC: *tu-sumbi* ***tw-ami*** *tu-tatu* (Martins 1990:163)
C10.SG-hen C10.AGR-1SG C10.AGR-three
hens my three
'my three hens'

**(b)** Possessive determiners always precede adjectives:

(90) CC: *ka-wa* ***k'ami*** *ka-kehe* (*ibid* 164)
C10.SG-dog C10.AGR-1SG C10.AGR-small
dog my small
'my small dog'

**(c)** Cardinal numerals always precede adjectives:

(91) CC: *a-tfu* ***a-wana*** *a-pema* (*ibid* 163)
C1.PL-person C1.AGR-four C1.AGR-good
men four good
'four good men'

### 4.2.2. Number marking strategies in KK, K, U and CC

Number marking in EP and KK, K, U and CC are alike to the extent that they are inflectional mechanisms and only allow for the establishment of singular-plural oppositions, but two fundamental differences between the number marking in the latter group of languages and that in EP have traditionally been identified in the literature. First, both singular and plural number must be formally marked by a single marker each, which is the same for every noun in a given noun class. Second, number is marked by means of prefixes, not suffixes. However, this traditional description of marking in KK, K, U and CC is simplistic to some extent. Hence, this section aims to tentatively describe the complexity of number marking in these languages so as to better evaluate the influence of the substrate in contemporary number marking strategies in AVP's noun phrase.

The first problem with the aforementioned description of number marking in KK, K, U and CC is that while both singular and plural number is indeed typically marked on most nouns, this needs not be so. In fact, as far as formal marking of number is concerned, I could identify four basic types: (1) marking of both singular and plural number (the most common type), (2) singular number marking only, (3) plural number marking only and (4) single form for both singular and plural number (i.e. singular-plural oppositions are expressed by means of the agreement markers on the non-nuclear elements in the NP)[22]. Most noun classes display only one type of number marking for all nouns, but some actually display two or more types[23]. This existence of variation concerning the types of nominals that display singular-plural oppositions, concerning the obligatory or optional use of these oppositions as well as their marking on the noun as opposed to agreement is not uncommon in the world's languages and appears to be related to constraints imposed by the Animacy Hierarchy (Corbett 2000:87-88):

**Figure 1: The Animacy Hierarchy (Corbett 2000:56)**

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Speaker | > | addressee | > | 3rd person | > | kin | > | human | > | animate | > | inanimate |
| (1st person pronouns) | | (2nd person pronouns) | | | | | | | | | | |

[22] Of course, the establishment of singular-plural oppositions exclusively by means of agreement markers is only possible because all agreement classes display a set of two distinct prefixes to mark singular and plural number in the targets, exception made to class 5 in CC which has the same prefix for both singular and plural number agreement (cf. tables 14, 15, 16 and 17).

[23] The data seems to indicate that the noun classes that allow for the use of different types of formal number marking are usually those which include nouns lower in the Animacy Hierarchy, e.g. borrowings (allow for type 3 and type 4 number marking), abstract entities (allow for type 2 and type 4 number marking), massive, collective, deverbal and irrational entities (allow for type 1 and one of two of the other types of marking).

According to Corbett (2000), the Animacy Hierarchy constrains number in three important ways:

I. The singular-plural distinction in a given language must affect a top segment of the Animacy Hierarchy (*ibid.* 56).

II. Lexical items may be irregular in terms of number marking with respect to the Animacy Hierarchy and regular in terms of agreement, but not vice-versa (*ibid.* 67).

III. As we move rightwards along the Animacy Hierarchy, the likelihood of number being distinguished will decrease monotonically (that is, with no intervening increase) (*ibid.* 70).

Tables 18, 19, 20 and 21 show the different types of number marking found in KK, K, U and CC, respectively. When the language does not display a given type of number marking this is identified by (---). Noun classes allowing for more than one type of number marking are identified in parenthesis.

**Table 18: Types of formal number marking in Kikongo (based on Tavares 1915)**

| STRATEGY | NOUN CLASS | EXAMPLES |
|---|---|---|
| 1. both SG and PL number is marked | 1 | **n**kentu 'woman' *vs.* **a**kentu 'women' |
| | (2) | **mu**anzi 'root, nerve' *vs.* **mi**anzi 'roots, nerves' |
| | (3) | **ki**nkutu 'shirt' *vs.* **in**kutu 'shirts' |
| | (4) | **di**nkonde 'banana' *vs.* **man**konde 'bananas' |
| | 5 | **ku**lu 'leg' *vs.* **ma**lu 'legs' |
| | 6 | **lu**muenu 'mirror' *vs.* **tu**muenu 'mirrors' |
| | 10 | **vu**ma 'place' *vs.* **mu**ma 'places' |
| 2. only SG number is marked | --- | --- |
| 3. only PL number is marked | (3) | lumbu 'day' *vs.* **i**lumbu 'days' |
| | (4) | *no data provided by Tavares (1915)*[24] |
| | (7) | sampatu 'shoe' *vs.* **zin**sapatu 'pair of shoes' |
| | 8 | ko 'father-in-law' *vs.* **ma**ko 'fathers-in-law' |
| 4. single form for both SG and PL number | (2) | *no data provided by Tavares (1915)*[25] |
| | (7) | *no data provided by Tavares (1915)*[26] |
| | 9 | uandu 'pea' and 'peas' |

[24] Tavares (1915:13) observes that "os nomes desta classe são, com frequência, empregados sem o prefixo *di* do singular".

[25] Tavares (1915:11) observes that "em muitos dialectos do *kikongo*, há actualmente uma grande tendência para o abandono do prefixo do plural dos nomes desta classe".

[26] Tavares (1918:17) observes that "há hoje uma enorme tendência para – em todos os dialectos do *kikongo* – se abandonar o prefixo plural dos nomes desta classe".

**Table 19: Types of formal number marking in Kimbundu (based on Chatelain 1888-1889)**

| STRATEGY | NOUN CLASS | EXAMPLES |
|---|---|---|
| 1. both SG and PL number is marked | 1 | **mu**hatu 'woman' *vs.* **a**hatu 'women' |
| | 2 | **mu**xi 'stick' *vs.* **mi**xi 'sticks' |
| | 3 | **ki**nama 'leg' *vs.* inama 'legs' |
| | (4) | **ri**bitu 'door' *vs.* **ma**bitu 'doors' |
| | 5 | **u**lungu 'canoe' *vs.* **mau**lungu 'canoes' |
| | 6 | **lu**muenu 'mirror' *vs.* **malu**muenu 'mirrors' |
| | 7 | **tu**bia 'fire' *vs.* **matu**bia 'fires' |
| | 8 | **ku**nua 'a drink' *vs.* **maku**nua 'drinks' |
| | 10 | **ka**hoji 'little lion' *vs.* **tu**hoji 'little lions' |
| 2. only SG number is marked | --- | --- |
| 3. only PL number is marked | (4) | *no data provided by Chatelain (1888-89)*[27] |
| | 9 | hoji 'lion' *vs.* **ji**hoji 'lions' |
| 4. single form for both SG and PL number | --- | --- |

**Table 20: Types of formal number marking in Umbundu (adapted from Nascimento 1893)**

| STRATEGY | NOUN CLASS | EXAMPLES |
|---|---|---|
| 1. both SG and PL number is marked | 1 | **omu**nu 'person' *vs.* **oma**nu 'people'<br>**u**kai 'woman' *vs.* **a**kai 'women' |
| | 2 | **omu**nda 'mountain' *vs.* **omi**nda 'mountains'<br>**u**ti 'tree' *vs.* **o**viti 'trees' |
| | 3 | **e**kapa 'potato' *vs.* **a**kapa 'potatos'<br>**i**su 'eye' *vs.* **ova**su 'eyes' |
| | 4 | **ochi**mbundu 'black man' *vs.* **ovi**mbundu 'black men |
| | 6 | **olu**si 'fish' *vs.* **olo**si 'fishes' |
| | 7 | **oka**kolototo 'dwarf' *vs.* **otu**kolototo 'dwarves' |
| | (8) | **ou**ato 'canoe' *vs.* **au**ato 'canoes' |
| | 9 | **oku**tui 'ear' *vs.* **ova**tui 'ears' |
| 2. only SG number is marked | (8) | **ou**elema 'darkness' |
| 3. only PL number is marked | 5 | ngombe 'ox' *vs.* **olo**ngombe 'oxen' |
| 4. single form for both SG and PL number | --- | --- |

[27] According to Chatelain (1888-1889:4), when a class 4 noun co-occurs with non-nuclear elements in the NP the singular marker is not used on the former but only on the latter, i.e. the class of the noun is identified by means of the agreement markers only.

**Table 21: Types of formal number marking in Cokwe (adapted from Martins 1990)**

| STRATEGY | NOUN CLASS | EXAMPLES |
|---|---|---|
| 1. both SG and PL number is marked | 1 | **mu**tfu 'person' *vs.* **a**tfu 'people' |
| | (2) | **mu**tondo 'tree' *vs.* **mi**tondo 'trees' |
| | 3 | **tchi**paya 'shoulder' *vs.* **i**paya 'shoulders' |
| | (4) | **ri**zo 'tooth' *vs.* **ma**zo 'teeth' |
| | | **ri**himbi 'small bee' *vs.* **mari**himbi 'small bees' |
| | 5 | **u**ta 'rifle' *vs.* **maw**ta 'rifles' |
| | 6 | **lu**nga 'man' *vs.* **malu**nga 'men' |
| | 7 | **tu**fulu 'thorax' *vs.* **matu**fulu 'thoraxes' |
| | 8 | **ku**mbi 'stork' *vs.* **maku**mbi 'storks' |
| | 10 | **ka**sumbi 'hen' *vs.* **tu**sumbi 'hens' |
| 2. only SG number is marked | (2) | **mu**njapela 'pocket' *vs.* njapela 'pockets' |
| 3. only PL number is marked | (9) | ngalavanta 'tie' *vs.* **ji**ngalavanta 'ties' |
| 4. single form for both SG and PL number | (2) | mwalwa 'the sun' |
| | (9) | yichi 'fish' |

The second problem with the traditional description of number marking in KK, K, U and CC is that while number is indeed marked mostly by means of prefixes, different strategies are used to add them to the noun. I could identify two main strategies: either (1) the singular marker is replaced with the corresponding plural marker, i.e. singular and plural markers are in complementary distribution or (2) the plural number marker is added in front of the singular number marker, i.e. the singular and plural markers coexist in the noun. These strategies are systematized in tables 4.12 to 4.15 below. Of course, they apply only to those noun classes that display singular-plural oppositions (i.e. type 1 marking).

**Table 22: Number marking strategies in Kikongo (based on Tavares 1915)**

| STRATEGY | NOUN CLASS | EXAMPLES |
|---|---|---|
| 1. SG marker replaced with PL marker | 1 | **n**kentu 'woman' *vs.* **a**kentu 'women' |
| | 2 | **mu**anzi 'root, nerve' *vs.* **mi**anzi 'roots, nerves' |
| | 3 | **ki**nkutu 'shirt' *vs.* **i**nkutu 'shirts' |
| | 4 | **di**nkonde 'banana' *vs.* **ma**nkonde 'bananas' |
| | 5 | **ku**lu 'leg' *vs.* **ma**lu 'legs' |
| | 6 | **lu**muenu 'mirror' *vs.* **tu**muenu 'mirrors' |
| | 10 | **vu**ma 'place' *vs.* **mu**ma 'places' |
| 2. PL marker added in front of SG marker | --- | --- |

**Table 23: Number marking strategies in Kimbundu (based on Chatelain 1888-1889)**

| STRATEGY | NOUN CLASS | EXAMPLES |
|---|---|---|
| 1. SG marker replaced with PL marker | 1 | **ka**hoji 'little lion' *vs.* **tu**hoji 'little lions' |
| | 2 | **mu**hatu 'woman' *vs.* **a**hatu 'women' |
| | 3 | **mu**xi 'stick' *vs.* **mi**xi 'sticks' |
| | 4 | **ki**nama 'leg' *vs.* **i**nama 'legs' |
| | 10 | **ri**bitu 'door' *vs.* **ma**bitu 'doors' |
| 2. PL marker added in front of SG marker | 5 | **u**lungu 'canoe' *vs.* **mau**lungu 'canoes' |
| | 6 | **lu**muenu 'mirror' *vs.* **malu**muenu 'mirrors' |
| | 7 | **tu**bia 'fire' *vs.* **matu**bia 'fires' |
| | 8 | **ku**nua 'a drink' *vs.* **maku**nua 'drinks' |

**Table 24: Number markering strategies in Umbundu (based on Nascimento 1893)**

| STRATEGY | NOUN CLASS | EXAMPLES |
|---|---|---|
| 1. SG marker replaced with PL marker | 1 | **omu**nu 'person' *vs.* **oma**nu 'people'<br>**u**kai 'woman' *vs.* **a**kai 'women' |
| | 2 | **omu**nda 'mountain' *vs.* **omi**nda 'mountains'<br>**u**ti 'tree' *vs.* **o**viti 'trees' |
| | 3 | **e**kapa 'potato' *vs.* **a**kapa 'potatos'<br>**i**su 'eye' *vs.* **o**vasu 'eyes' |
| | 4 | **ochi**mbundu 'black man' *vs.* **ovi**mbundu 'black men' |
| | 6 | **olu**si 'fish' *vs.* **olo**si 'fishes' |
| | 7 | **oka**kolototo 'dwarf' *vs.* **otu**kolototo 'dwarves' |
| | 8 | **ou**ato 'canoe' *vs.* **au**ato 'canoes' |
| | 9 | **oku**tui 'ear' *vs.* **ova**tui 'ears' |
| 2. PL marker added in front of SG marker | --- | --- |

**Table 25: Number marking strategies in Cokwe (based on Martins 1990)**

| STRATEGY | NOUN CLASS | EXAMPLES |
|---|---|---|
| 1. SG marker replaced with PL marker | 1 | **mu**tfu 'person' *vs.* **a**tfu 'people' |
| | 2 | **mu**tondo 'tree' *vs.* **mi**tondo 'trees' |
| | 3 | **tchi**paya 'shoulder' *vs.* **i**paya 'shoulders' |
| | (4) | **ri**zo 'tooth' *vs.* **ma**zo 'teeth' |
| | 10 | **ka**sumbi 'hen' *vs.* **tu**sumbi 'hens' |
| 2. PL marker added in front of SG marker | (4) | **ri**himbi 'small bee' *vs.* **mari**himbi 'small bees' |
| | 5 | **u**ta 'rifle' *vs.* **mawta** 'rifles' |
| | 6 | **lu**nga 'man' *vs.* **malu**nga 'men' |
| | 7 | **tu**fulu 'thorax' *vs.* **matu**fulu 'thoraxes' |
| | 8 | **ku**mbi 'stork' *vs.* **maku**mbi 'storks' |

Hence, as far as number marking in KK, K, U and CC is concerned, the tables above are enlightening in two important ways. First, they show that while the marking of both singular and plural number is indeed the most common type of number marking in these languages, in a considerable number of their noun classes only either singular (i.e. type 2) or plural (i.e. type 3) number is formally marked on the noun. Moreover, these tables show that, also unlike what is traditionally argued in Bantuist literature, number can, and often is, marked on the non-nuclear elements of the NP but not on the head noun (i.e. type 4), which is particularly interesting when compared to contemporary number marking strategies in AVP (cf. section 4.3.2 below).

### 4.2.3. Gender marking in KK, K, U and CC

The grammatical category of gender in the Bantu languages differs from that of EP in two important ways. First, in these languages, gender contrasts transcend the sex-based gender oppositions that are typical of most Indo-European languages: feminine vs. masculine vs. neuter. Moreover, sex-gender contrasts are usually only displayed by [+ animate] nouns. Second, sex-based gender contrasts in the Bantu languages are not morphological but lexical, i.e. different words are added to the head noun (but not to the non-nuclear elements in the NP) depending on whether the noun refers to a male, female or neuter entity. The addition of these words is not required with those [+ animate] nouns which are inherently either male, female or neuter (e.g. kinship terms, some animal names, etc).

Despite these general tendencies, sex-based gender contrasts in Bantu also display some language-specific features. For example, while animacy is the key feature in determining which nouns can display sex-based gender contrasts, different words can be used to express these contrasts depending on semantic properties of the noun other than animacy, e.g. [+ human] *vs.* [- human], [+ rational] vs. [- rational], [+ definite] vs. [- definite], etc. The sex-based gender values displayed by each language differ significantly also (e.g. some languages allow only for masculine-feminine oppositions while others allow for masculine-feminine-neuter contrasts). The word order in these contrasts is also language-specific (e.g. in some languages, such as KK, the word expressing the sex-gender contrasts must follow the noun, but in other languages, such as Cokwe, the word order depends on the semantic properties of the noun). Moreover, class agreement with the head noun may be obligatory or optional depending on the language.

Therefore, this section aims to tentatively describe the sex-based gender systems of KK, K, U and CC concerning four fundamental features: (1) the nature of the semantic traits that determine which nouns can or cannot display sex-based gender contrasts, (2) the sex-gender values displayed, (3) the word order, and (4) the existence or absence of class agreement with

the head noun. Hopefully, this will lead to a more comprehensive understanding of gender marking and agreement in AVP.

In Kikongo, only nouns denoting [+ animate] entities can express sex-based gender oppositions, which in this language are limited to two values only, i.e. feminine and masculine. The feminine is expressed by the addition of the adjective *a-nkentu* to the noun whose sex-gender one wants to specify, but two different adjectives are used to express the masculine depending on whether the noun denotes an entity that is [+ human] or [- rational]. These two adjectives are *a-kala* and *a-mbakala*[28], respectively[29]. Hence, animacy alone determines whether a noun can enter feminine-masculine oppositions, but whereas the form of the feminine is not constrained by any other semantic feature, as in (92) and (93), the form of the masculine depends on the humanness and rationality of the entity that is denoted by the noun, as shown in (94) and (95). Concerning the word order, the available data shows that the adjectival stems that mark feminine-masculine oppositions must always follow the head noun and agree in class and number with it[30].

(92) KK: *mu-an'* ***a-nkentu***
C1-child ADJ.CLF-female
'girl'

(93) KK: *di-mene* ***di-a-nkentu***
C4-sheep C4.AGR-ADJ.CLF-female
'female sheep'

(94) KK: *di-mene* ***di-a-mbakala***
C4-sheep C4.AGR-ADJ.CLF-male
'male sheep'

(95) KK: *mu-an'* ***a-kala***
C1-child ADJ.CLF-male
'boy'

In Kimbundu, as in KK, only nouns denoting [+ animate] entities can express sex-based gender oppositions. The sex-gender values allowed are also only feminine and masculine. However, unlike KK, feminine-masculine contrasts in K are expressed by means of nominal stems, not adjectives. These are *-hetu* or *-ngambo* to express the feminine and *-ndata* to express the masculine. Both must follow the head noun, as shown in the examples below drawn from Maia (1957:10):

[28] Tavares (1915:25) observes that "Em alguns dialectos do *kikongo*, emprega-se *a-koko* com a mesma significação que *a-mbakala*".

[29] Notice that the prefix *a-* before each of these adjectives is not a class marker but a derivational prefix (i.e. derives adjectives from verbs). This prefix is always the same regardless of the class agreement marker that precedes it.

[30] Examples (92) and (95) seem to suggest that there is no class agreement, but this is because the class 1 agreement marker is omitted to avoid three consecutive vowels (i.e. *mu-ana* ***u-a-kala*** or *mu-ana* ***u-a-nkentu***). In (93) and (94) the class 4 agreement marker is not omitted because it is constituted by a consonant and vowel (i.e. *di-*).

(96) K: *o* *mo-na* *o* ***mu-ndata***
the C1-child the C1.AGR-man
'son'

(97) K: *o* *mo-na* *o* ***mu-hetu***
the C1-child the C1.AGR-woman
'daughter'

(98) K: *o-mbua* *io* ***mu-ndata***
C6-dog the C1.AGR-man
'male dog'

(99) K: *o-mbua* *io* ***mu-hetu***
C6-dog the C1.AGR-woman
'female dog'

Notice that in (96) and (97) the agreement markers match the class of the head noun (i.e. class 1), but in (98) and (99) there is a mismatch between the noun class marked on the head noun itself and the agreement class it is assigned to (i.e. the head noun is class 6 whereas the agreement markers are class 1). As in (80) above, this is a clear example of quirky agreement. The reason for this is that Maia (1957:10) takes the morphological form of the noun as the key criterion to determine its class and assigns *mbua* to class 6. However, the agreement markers on both *ndata* and *hetu* show that, regardless of its morphological form, *mbua* is actually a class 1 noun. In fact, the class 6 identified by Maia (1957) does not match any of the classes identified by Chatelain (1888-1889). This seems to validate the hypothesis that this mismatch results from a mistaken analysis of the data, although this does not invalidate the fact that, in K, the nouns that specify the sex-gender values must indeed agree in class and number with the head noun.

In Umbundu there is also no morphological distinction between nouns denoting female entities and those denoting male ones. Hence, the sex-gender values of feminine and masculine of nouns denoting [+ animate] entities are expressed by means of the adjectival stems *–kai* and *–lumes*, respectively. These stems are preceded by the marker that agrees in class and number with the head noun whose sex-gender one wants to specify. There is no neuter sex-gender value in Umbundu. Hence:

(100) U: *o-njila* ***o-nkai*** (Nascimento 1893:20)
C5-bird C5.AGR-female
'female bird'

(101) U: *o-njila* ***o-nlume*** (*ibid.*)
C5-bird C5.AGR-male
'male bird'

Notice that (100) and (101) refer only to nouns that denote [- human] entities. In fact, Nascimento (1893) does not provide any examples for nouns denoting [+ human] entities. He only observes that there is a limited number of nouns "com significação *genérica especial*"

(*ibid.* 21), i.e. some nouns have different forms depending on whether they refer to female or male entities. This group of nouns includes both [+ human] and [-human] nouns (e.g. *u-lume* 'man' vs. *u-kai* 'woman', *oku-enju* 'boy' vs. *u-feku* 'girl', *o-ngombe* 'steer' vs. *o-njindi* 'cow'). Hence, it is not possible to determine whether sex-gender oppositions are established in different ways for nouns denoting [+ human] as opposed to [- human] entities. It is possible to confirm though that *–kai* (F) and *–lume* (M) must always follow the noun.

As far as Cokwe is concerned, sex-based gender oppositions are also expressed by means of adjectival stems. However, the system of sex-gender marking in Cokwe is more complex than that of KK, K and U. First, Cokwe distinguishes between three sex-gender values (i.e. feminine, masculine and neuter). Second, both nouns denoting [+ animate] and nouns denoting [- animate] entities can contrast in sex-gender. Third, these contrasts are not expressed in the same way in all nouns denoting [+ animate] entities. This is because these nouns are not only marked for different sex-gender values (e.g. feminine-masculine vs. feminine-masculine-neuter) but they also take different sex-gender markers depending on whether they refer to birds or to humans and personified entities. Finally, as far as word order is concerned, sex-gender markers may occur in two positions: (a) after the noun or (b) either before or after the noun. The word order too is constrained by the semantic properties of the noun.

Hence, the system of sex-based gender marking in Cokwe is presented in table 26 below. Notice that the order in which the semantic features of the head noun are arranged is important (i.e. things, birds, humans and personified things, animate entities) because if animates came first in the table then both nouns referring to birds and people (as well as personified things) would take the sex-gender markers *tchihwo* (F) and *kunji*.

**Table 26: Sex-based gender oppositions in Cokwe (based on Martins 1990:33-34)**

| SEMANTICS | SEX-GENDER | EXAMPLES | WORD ORDER |
|---|---|---|---|
| **Things** | *Neuter* | (102)***tchuma*** 'thing' ~ ***yuma*** 'things'<br>(103)***imate*** 'thing' ~ ***yamate*** 'things' | After the noun |
| **Birds** | *Masculine* | (104)*ka-jia* *wa* ***ndemba***<br>C10-bird Gen male<br>'male bird'<br>(105)***ndemba*** *ka-jia*<br>male C10-bird<br>'male bird' | Before *or* after the noun |
| | *Feminine* | (106)*ka-jia* *wa* ***tchari***<br>C10-bird GEN female<br>'female bird'<br>(107)***tchari*** *ka-jia*<br>female C10-bird<br>'female bird' | |
| **Human (personified things)** | *Masculine* | (108)*mu-tfu* *wa* ***lunga***<br>C1-person GEN male<br>'man' or 'male person' | After the head noun |
| | *Feminine* | (109)*mu-tfu* *wa* ***pfwo***<br>C1-person GEN female<br>'woman' or 'female person' | |
| | *Neuter* | (110)*mu-tfu* ***kachi-ku***<br>C1- person N-INDF<br>'not a single person' or 'there is nobody' | |
| **Animate** | *Masculine* | (111)***kunj'*** *a* *panga*<br>male GEN sheep<br>'male sheep' or 'sheep of a male'<br>(112)*panga* *wa* ***kunji***<br>sheep GEN male<br>'male sheep' or 'sheep of a male' | Before *or* after the noun |
| | *Feminine* | (113)***tchihwo*** *tcha* *panga*<br>Female GEN sheep<br>'female sheep' or 'sheep of a female'<br>(114)*panga* *wa* ***tchihwo***<br>sheep GEN female<br>'female sheep' or 'sheep of a female' | |

### 4.2.4. Case and word order of person markers in KK, K, U and CC

KK, K, U and CC display two major sets of person markers: those denoting [+ human] entities (which typically belong to class 1) and those denoting [- human] entities (which belong to any noun class other than class 1). The former distinguish between first, second and third persons, whereas the latter typically only mark the third person. Nonetheless, both [+ human] and [-human] person markers contrast in number (i.e. singular vs. plural) and case (i.e. subject vs.

object), but not in gender. Despite these similarities between the four languages, there are also language-specific features, especially concerning the relation between the function of person markers (i.e. subject vs. object) and their morphophonological form (i.e. free vs. bound)[31]. Nonetheless, there is a generalization one can make concerning this issue: KK, K, U and CC all have a set of free subject markers denoting [+ human] entities which can replace the head noun in the NP. These are shown in table 27[32]:

**Table 27: Free subject markers denoting [+ human] entities in KK, K, U and CC[33]**

| | KIKONGO (Tavares 1915:34) | KIMBUNDU (Chatelain 1888:154) | UMBUNDU (Nascimento 1893:35) | COKWE (Martins 1990:50) | |
|---|---|---|---|---|---|
| 1SG | *mono* | *eme* | *ame* | *yami* | 'I' |
| 2SG | *nge, ngei* | *Eie* | *ove* | *Yena* | 'you' |
| 3SG | *iandi* | *muene* | *eie* | *yie, mwene* | 'he/she, it' |
| 1PL | *ietu* | *Etu* | *etu* | *yetwe* | 'we' |
| 2PL | *ienu* | *Enu* | *ene* | *yenuwe* | 'you' |
| 3PL | *iau* | *Ene* | *ovo* | *ene, yo, ayo* | 'they' |

Object person markers denoting [+ human] entities in KK, K, U and CC display different morphophonological forms. On the one hand, in KK object markers are free forms, but in K, U and CC they are exclusively bound forms (i.e. infixes in K and U; suffixes or infixes in CC depending on whether they function as direct or indirect object markers). As a consequence of their different morphophonological forms, person markers in these four languages also have different word orders. Hence, in K object person markers always precede the verb form, i.e. they are infixed between the subject agreement marker and the stem of the verb, as in (115), but in Cokwe, while indirect object markers always precede the verb (i.e. as infixes), direct object markers must follow it (i.e. as suffixes), as in (116).

[31] It is important to notice that the term bound is not used here as "as a cover term for both person affixes and clitics" (Siewierska 2004:24); rather "it designates only person markers expressed by affixes or much less frequently via changes to the stem" (*ibid.*). The distinction between clitic and bound forms is yet another reason for preferring the term person marker to personal pronoun. In fact, the bound forms in KK, K, U and CC, as the suffixes used in EP verbs to mark person and number, are undoubtedly person markers, but they can hardly be said to match the traditional definition of pronoun

[32] These person markers are referred to in the grammars surveyed here as 'pronomes pessoais primários' or 'pronomes pessoais absolutos'.

[33] In all the four languages they are reflected in the verb form by means of prefixes which agree in person and number with them. When the subject of the verb is not a free person marker but a noun belonging to a noun class other than class 1, the prefix on the verb agrees in class and number with the head noun. The bound person markers are extremely important as they are often the sole expression of the subject in a sentence. However, since they are attached to the verb phrase, and not the NP, they will not be analyzed in detail here.

(115) K: *eie* *u<ngi>zola* (Chatelain 1888-1889:73)
you 2SG.AGR<me>PRS.love
'you love me'

(116) CC: *ka-wa* *yami* *na<mu>hele-yo* (Martins 1990:165)
C10.SG-dog I AUX<3SG>give-3PL
'the dog, I gave it to him '

The different object person markers denoting [+ human] entities in KK, K, U and CC are displayed in table 28 below:

**Table 28: Object person markers denoting [+ human] entities in KK, K, U and CC**

| | KIKONGO (Tavares 1915:34) | KIMBUNDU (Chatelain 1888:154) | UMBUNDU (Nascimento 1893:35) | COKWE (Martins 1990:50) | |
|---|---|---|---|---|---|
| 1SG | *Kuame* | *-ngi-* | *-ndi-* or *–ngu-* | *-ngu-* | 'me' |
| 2SG | *Kuaku* | *-ku-* | *-ku-* | *-ku-* | 'you' |
| 3SG | *Kuandi* | *-mu-* | *-u-* or *–o-* | *-ka-* or *–mu-* | 'him, her' |
| 1PL | *Kuetu* | *-tu-* | *-tu-* | *-tu-* | 'us' |
| 2PL | *Kuenu* | *-nu-* | *-vu-* | *-nu-* | 'you' |
| 3PL | *Kuau* | *-a-* | *-va-* | *-ka-* or *–a-* | 'them' |

KK, K, U and CC also display reflexive and reciprocal person markers (cf. table 29) as well as emphatic person markers (cf. table 30) to denote [+ human] entities.

**Table 29: Reflexive person markers denoting [+ human] entities in KK, K, U and CC**

| | KIKONGO (Tavares 1915:34) | KIMBUNDU (Chatelain 1888:81) | UMBUNDU (Nascimento 1893:50) | COKWE (Martins 1990:51) |
|---|---|---|---|---|
| 1SG | *kuandi* | *-ri-* | *-li-* | *-ri-* |
| 2SG | | | | |
| 3SG | | | | |
| 1PL | *kuau* | | | |
| 2PL | | | | |
| 3PL | | | | |

Notice that, unlike EP, there is typically a single reflexive or reciprocal person marker for all persons and numbers. KK is the only language that displays two different reflexive person markers (i.e. singular vs. plural), but Tavares (1915) does not provide sufficient data to account for this.

**Table 30: Emphatic person markers denoting [+ human] entities in KK, K, U and CC**

| | KIKONGO (Tavares 1915:34) | KIMBUNDU (Chatelain 1888:81) | UMBUNDU (Nascimento 1893:50) | COKWE (Martins 1990:51) | |
|---|---|---|---|---|---|
| 1SG | *mono kuame* | *eme muene* | | | 'myself' |
| 2SG | *nge kuaku* | *eie muene* | | | 'yourself' |
| 3SG | *iandi kuandi* | *muene muene* | [34] | [35] | 'himself' |
| 1PL | *ietu kuetu* | *etu muene* | | | 'ourselves' |
| 2PL | *ienu kuenu* | *enu muene* | | | 'yourselves' |
| 3PL | *iau kuau* | *ene muene* | | | 'themselves' |

When the person markers refer to nouns denoting [- human] entities, which in Bantu belong to any noun class other than class 1, they are preceded by the corresponding class prefix, hence agreeing in class and number with the head it replaces (cf. tables 14 to 17 for a list of the agreement markers in KK, K, U and CC). These person markers also contrast in case, hence establishing an opposition between subject (A) and object (B) person markers. This is illustrated below with examples from KK:

**(A)**

(117) KK: *e meme di-a-mbakala kuevi d'in'e?*
the sheep C4.AGR.ADJ.CLS-male where is?
'Where is the male sheep?'

(118) KK: ***di-au*** *d'ina vava*
C4.AGR.it is here
'It is here.'

**(B)**

(119) KK: *u-mon' e meme di-a-mbakal'e?*
you-see the sheep C4.AGR.ADJ.CLS-male
'Do you see the sheep?'

(120) KK: *i-mona* ***di-o***
1SG.COP- I C4.AGR-it
'I see it.'

[34] No data on emphatic person markers is provided in Nascimento (1893).

[35] According to Martins (1990:51) the co-occurrence of the free subject markers with the subject agreement markers on the verb is itself emphatic:

> O pronome pessoal, no caso nominativo, a não ser que seja seguido do verbo *ku-ri* ou *ku-pwa*, só se emprega enfaticamente, pois nos restantes verbos o prefixo verbal, ou verbo pronominal, já designam por si mesmos o sujeito, ou então para reforçar o emprego do sujeito.

### 4.2.5. Possession in KK, K, U and CC

KK, K, U and CC express possession mostly by means of possessive determiners, which are shown in table 31 below[36]:

**Table 31: Possessive determiners in KK, K, U and CC**

| | KIKONGO (Tavares 1915:38) | KIMBUNDU (Chatelain 1888-9:15) | UMBUNDU (Nascimento 1893: 39) | COKWE (Martins 1990:53) | |
|---|---|---|---|---|---|
| 1SG | *ame* | *ami* | *ame* | *ami* | 'my, mine' |
| 2SG | *aku* | *é* | *ove* | *we* | 'your, yours' |
| 3SG | *andi* | *ê* | *eie* | *enye* | 'his, her, hers, its' |
| 1PL | *etu* | *etu* | *etu* | *etwe* | 'our, ours' |
| 2PL | *enu* | *enu* | *ene* | *enwe* | 'your, yours' |
| 3PL | *au* | *â* | *ovo* | *wo* | 'their, theirs' |

As shown in the table above, possessive determiners in KK, K, U and CC display different forms depending on whether they refer to first, second or third person possessors, both singular and plural. Moreover, they must agree in class and number with the possessed noun by means of the agreement markers presented in tables 14 to 17. Possessive determiners always follow the possessed noun. Hence:

(121) KK: *tual'* *e* *meme* ***di-ame*** *i* *e* ***di-aku***
2SG.IMP.bring the sheep C4.AGR-1SG.POSS and DEF C4.AGR-2SG.POSS
'bring my sheep and yours' (Tavares 1915:39)

(122) K: ***mu-bika*** ***u-ami***
C1.SG-slave C1.AGR-1SG.POSS
'my slave' (Chatelain 1888-1889:15)

(123) U: *e-timba* ***li-eie***
C3.SG-body C3.AGR-3SG.POSS
'his body' (Valente 1964:132)

(124) CC: ***mu-tondo*** ***w-enye***
C2.SG-tree C2.AGR-3SG.POSS
'his/her tree' (Martins 1990:53)

[36] These are traditionally referred to in the literature as possessive pronouns, but Mutaka (2000:194) notes the inaccuracy of this designation in reference to Bantu possessive pronouns:

> We prefer to refer to it as a possessive determiner because it functions in limiting the potential domain of reference of a relevant noun, just like any other determiner, and because unlike other pronouns, the possessive usually does not function independently as a noun phrase.

Possession can also be expressed by means of the genitive, which is realized by a preposition whose meaning is similar to the EP preposition *de*. Table 32 shows the genitive particles used in KK, K, U and CC:

**Table 32: Genitive particles in KK, K, U and CC**

| KIKONGO (Tavares 1915:101) | KIMBUNDU (Chatelain 1888:10) | UMBUNDU (Nascimento 1893:40) | CHOKWE (Martins 1990:50) | |
|---|---|---|---|---|
| *a* | *a* | *o* | *a* | 'of' |

The genitive particle is always the same, regardless of the noun class of the possessed noun, but it is typically preceded by the corresponding class-number agreement marker (cf. tables 14 to 17 for a list of the agreement markers in KK, K, U and CC). As far as the word order is concerned, the genitive must also follow the possessed noun. This is illustrated with the examples from KK and K below:

(125) KK: *e zi-nzo z-a n-tinu* (Tavares 1915:101)
the C7.PL-house C7.PL-GEN C2.SG-king
'the houses of the king' or 'the king's houses'

(126) K: *ji-nzo j-a mu-ndele* (Chatelain 1888-89:14)
C9.Pl-house C9.AGR-GEN C1.SG-white man
'the houses of the white man' or 'the white man's house'

## 4.3. The noun phrase in Angolan Vernacular Portuguese

### 4.3.1. Internal structure and word order

AVP does not differ from EP regarding the number of elements in the NP or their grammatical categories. However, it does display significant differences regarding the word order of those elements, especially possessives and person markers. Number and gender marking and agreement also obey patterns that are essentially different from those used in EP. The following sections provide a description of these differences as well as a first set of explanations to account for them.

### 4.3.2. Number marking and agreement

As shown in section 4.1.2, EP requires that all the variable determiners, adjectives and quantifiers as well as the head of the NP be marked for number (Mira Mateus *et al.* 2003:330) by adding the plural number marker *–s*[37], as in (127):

[37] For an opposing view see Naro & Scherre (2000), who claim that "despite the monolithic anti-variation front we encountered among our fellow linguists in Portugal, there are some exceptions in the literature. The most

(127) EP: Todo*s* estes livros novos são importantes. 'All these new books are important.'

However, several linguists have shown that in AVP the head of the NP is seldom marked for number, as plurality is often indicated by adding *–s* only to the leftmost elements in the NP.

| | | | | |
|---|---|---|---|---|
| (128) | AVP: | *tem muitas mulher_* | 'he has many women' | (Mendes 1985:149) |
| | EP: | tem muitas mulheres | | |
| (129) | AVP: | *vigia as criança_* | 'watch over the children' | (Mingas 2000:67) |
| | EP: | vigia as crianças | | |
| (130) | AVP: | *estas duas mulher_* | 'these two women' | (Cabral 2005:75) |
| | EP: | estas duas mulheres | | |

Consequently, the non-nuclear elements in the NP seldom agree in plural number with the head noun. In fact, Cabral (2005:73), building on the analysis of 135 written documents produced by Angolan students (of whom 60 attend primary school, 50 attend secondary school and 23 were applying to a university), shows that the lack of agreement (both nominal and verbal) is not only the most common morphosyntactic feature of AVP in his corpus (i.e. 32%) but also the one whose incidence is apparently not constrained by formal instruction (i.e. 53 instances in primary school students, 40 in secondary school students and 53 instances in university applicants).

I also found this feature of AVP both in the speech of older and less educated speakers, as in (131) and (132), and in the informal speech of those who are younger and more educated, as in (133) and (134):

| | | | |
|---|---|---|---|
| (131) | AVP: | *as folha_ é verde* | 'the leafs are green' |
| | EP: | as folhas são verdes. | |
| (132) | AVP: | *as mulher_ é seis* | 'there are six women' |
| | EP: | as mulheres são seis | |

notable is a monograph entitled *A linguagem dos pescadores da Ericeira* by Joana Lopes Alves, first published in Lisbon in 1965" (*ibid.* 239). Moreover, Naro & Scherre (2000) argue that "Texts from earlier stages of Portuguese, particularly those from the period before grammatical norms were established during the classical stage of the sixteenth century, also occasionally show lack of agreement markers" (*ibid.* 242). Hence, Naro & Scherre (2000) declare themselves under "the impression that variable agreement can be found throughout the entire territory of Portugal, both within the noun phrase and between the verb and the subject" (*ibid.* 241). This hypothesis is interesting, but it strikes us as premature, not only because evidence of variable agreement is "statistically rare" (*ibid.*), but also because the occurrence of this phenomenon in most of the data provided by Naro & Scherre (2000) can be accounted for by different semantic and pragmatic explanations (e.g. *foi há muito ano* 'that was many years ago' and *andei por muito sítio* 'I've been to many places' are indeed common sentences in EP, but these seem to be best interpreted as emphatic set expressions). Consequently, further studies are needed before one can state beyond doubt that variable agreement is in fact a structural feature of EP rather than the result of performance errors (cf. the definition of error given on section 4.1). The fact that "the current state of the art in linguistic theory" cannot specify at what "input level there is a switch from performance to competence" (*ibid.*) cannot be used as an argument in favour of the validity of Naro & Scherre's hypothesis.

(133) AVP: *saber ler e escrever as carta_* 'be able to read and write the letters'
EP: saber ler e escrever as cartas

(134) AVP: *ajudo os papai_ em casa.* 'I help my parents at home'
EP: ajudo os pais em casa

For Marques (1983:219), Gärtner (1989:159) and Vilela (1999:56), the lack of number marking on the head of the NP is a consequence of the fact that in the Bantu languages number is marked on nouns by means of prefixes rather than suffixes, as in (133) to (134) above. Hence, the aforementioned linguists maintain that, in acquiring Portuguese, Angolans interpret the Portuguese article as equivalent to the Bantu prefix that typically marks class-number on the head noun.

Marques (1983), Gärtner (1989) and Vilela (1999) are correct in considering that number marking strategies in the NP in contemporary AVP are highly influenced by the structure of the substrate languages, but their description of this influence does not seem to be accurate as it can only account for instances in which the NP is composed of an article and a head noun, as in (131) to (134) above, but not for instances such as (135) to (137), in which the head noun is preceded by non-nuclear elements that belong to grammatical categories other than articles:

(135) AVP: *é mãe de três filho_* 'I am a mother of three'
EP: sou mãe de três filhos.

(136) AVP: *ele marca muitos golo_* 'he scores many goals'
EP: ele marca muitos golos.

(137) AVP: *aqueles prédio_ que estão ali* 'those buildings over there'
EP: aqueles prédios que estão ali

Moreover, when the head noun co-occurs with a definite article and a possessive determiner, both are marked for number[38]:

(138) AVP: *os meus filho?* 'my children?'
EP: os meus filhos

(139) AVP: *os meus passatempo_* 'my hobbies'
EP: os meus passatempos

(140) AVP: *se os meus pai_ tiverem dinheiro* 'if my parents have the money'
EP: se os meus pais tiverem dinheiro

(141) AVP: *os meus estudo_* 'my studies'
EP: os meus estudos

[38] Further data are needed to evaluate whether this applies to all NPs containing more than two non-nuclear elements (e.g. demonstrative + adjective; demonstrative + indefinite; etc) or just to NPs with definite articles and possessive determiners.

Hence, when the NP consists either of a head noun and one non-nuclear element[39] or of a head noun and two non-nuclear elements (especially if determiners and quantifiers), number tends to be exclusively marked by means of agreement markers on the modifiers but not on the head of the NP. This hypothesis is reinforced by the fact that the *–s* disappears even in some nouns in which it is not a plural number marker, as in (143):

(142) AVP: *Quando acabar as féria_* 'when the holidays are over'
EP: Quando acabarem as ferias.

The lack of plural number marking on the head of the NP can be accounted for by two complementary constraints imposed by the structure of the Bantu languages which were in closer contact with Portuguese in Angola. On the one hand, there is evidence in the literature that the Portuguese plural marker *–s* was not interpreted as such by Angolans learning Portuguese as a second language (cf. discussion of Oliveira 1990:89 in chapter 2, section 2.1.2). In fact, the marker was systematically deleted in Portuguese loans to the aforementioned languages regardless of their endings. In loans ending in consonant, whose plural marker in EP is *–es,* only the vowel [i] was maintained, probably as a means of fitting the loan to the typical structure of Bantu syllables, which are always open. The plural was marked by means of Bantu class prefixes. Chatelain's (1894:255) quote below exemplifies the treatment given to such Portuguese loans in Kimbundu:

> As the Portuguese were the first whites with whom the Angolans came in contact, and as the natives take at first all whites to be kinsmen, the name *Putu* was extended to all "white man's land" [...] Thus my native lad from Malanje called America *Putu ia ngeleji*, i.e., "the Putu of the English". In Angola, when a white man is found not to be a Portuguese, he is called a *ki-ngeleji,* pl. *I-ngeleji*, from the Portuguese "Inglez" [*author's emphasis*].

It is interesting to notice that this phonological rule seems to be active still in contemporary AVP (at least in the variety spoken in Dundo), as I was able to find utterances in which the plural marker *–s* disappears but the vowel [i] is maintained. Example (143) was produced by a 17-year-old woman, and (144) was uttered by an old man whose age I was unable to confirm:

(143) AVP: *outro bébé só tem dez mesi* 'the other baby is only 10 months old'
EP: o outro bébé só tem dez meses

[39] When the non-nuclear element is a numeral, the lack of number marking on the head noun is not a consequence of the number of elements in the NP. It is rather one of the four possibilities for marking plurals with numerals. In fact, Rijkhoff (2002:29) identifies four types of quantification strategies in the world's languages: (1) numeral + noun + plural (no classifier); (2) numeral + noun (plural marking on the noun, no classifier); (3) numeral + classifier + noun (no plural) and (4) numeral + classifier + noun + plural. "There are very few (possibly no) languages of type 4" (*ibid.*), but type 1, type 2 and type 3 quantification strategies are attested in Dutch, Oromo (a Cushitic language) and Thai, respectively.

(144) AVP: *ante portuguesi irem embora* 'before the Portuguese went away'
EP: antes de os portugueses irem embora

On the other hand, as shown in section 4.2.2, the data from 19th and early 20th century grammars of KK, K, U and CC indicate that in a considerable number of noun classes the plural number markers were often not used, so plural number was exclusively marked by means of agreement markers on the non-nuclear elements in the NP. While similar phenomena of lack of plural number agreement are also found in Archaic Portuguese texts, this can hardly account for these phenomena in contemporary AVP, not only because instances of variable plural agreement in Archaic Portuguese are 'statistically rare' (Naro & Scherre 2000:241), but also because the use of Portuguese was not widespread in Angola until the late 19th century (i.e. in the urban centers along the coast) and mid-twentieth century (i.e. in the interior, especially in the eastern provinces of the north and south of the country). Hence, the lack of plural number marking on the head noun seems to be best interpreted as a result of a latent tendency in the Bantu languages to mark number in some noun classes exclusively by means of agreement[40], which probably reinforced the previous use of a phonological rule of *–s* deletion in Portuguese loans to Bantu.

Nonetheless, the fact that number in the Bantu languages is marked to the left of the noun by means of prefixes must also be considered a plausible explanation to account for number marking strategies in AVP (Marques 1983; Vilela 1999). In fact, leftmost elements do tend to attract plural number marking, regardless of whether they are the first element in a compound noun, as in (146) and (147), or the head of the NP, as in (145):

(145) AVP: *há coisas que eles próprio exigem*
'there are things they themselves demand'
EP: há coisas que eles próprios exigem

(146) AVP: *afros-americano_* 'Afro-Americans'
EP: afro-americanos

(147) AVP: *guardas-chuva* 'umbrellas'
EP: guarda-chuvas

In fact, it is the attraction of plural number marking by the leftmost elements in the NP that seems to account for those instances in which the head noun in AVP is exceptionally marked for number, as in (148) to (151):

[40] This hypothesis, if tested with data from BVP, can offer a view of the origin of similar phenomena of variable number agreement in BVP other than the one presented by Naro & Scherre (2000), i.e. a European origin.

(148) AVP: *em Angola temos muitas línguas materna_*
'there are many mother tongues in Angola'
EP: em Angola temos muitas línguas maternas

(149) AVP: *é uma das línguas mais falada_ no mundo*
'[Portuguese] is one of the most widely spoken languages in the world'
EP: é uma das línguas mais faladas no mundo

(150) AVP: *... os dôs grandes período são...* (Cabral 2005a:23)
'the two main periods'
EP: os dois grandes períodos

(151) AVP: *conheço outras ciências auxiliar* (Cabral 2005a:31)
'I know other auxiliary sciences'
EP: conheço outras ciências auxiliaries

In short, our analysis of the data above enables us to hypothesize that in AVP plural number is always marked on the leftmost elements in the NP, be they heads or non-nuclear elements. This general rule is then specified according to the number of elements in the NP as follows: (a) when the NP consists of two elements, only the leftmost one is marked for number; and (b) when the noun phrase is composed of three elements, only the two first elements receive plural marking. The grammatical category of the elements in the NP and the syntactic function of the NP itself do not seem to constrain number marking.

### 4.3.3. Gender marking and agreement

As shown in section 4.3.2, linguists are agreed on the existence of variable number marking and agreement in AVP's noun phrase as a consequence of its tendency to mark plural number only on leftmost elements. An analysis of the available data allows us to identify stable patterns of number marking which can be traced back to the structure of the Bantu languages that were in closer contact with Portuguese in Angola, although further data are needed to sustain more definite conclusions. However, as far as sex-gender marking and agreement are concerned, there is very little available information in the literature. On the one hand, linguists' remarks on this topic are often too brief and simplistic:

> Sendo a distinção sexual inexistente em Kimbundu, os locutores da variante [i.e. Luanda Portuguese] apresentam uma incapacidade em fazer a diferença entre o acordo do determinante com um nome do género masculino e/ou feminino ... (Mingas 2000:83)

> No quimbundo o número recebe tantas variações quantas as classes (...) Por darem menos importância ao género aparecem no seu discurso discordâncias como "no mesmo barriga" (Mendes 1985:214)

On the other hand, because there are so few linguists working on AVP and because they tend to produce descriptions of the variety that are too general rather than detailed analyses of specific

linguistic phenomena, the data are few and most often repeated in publications by different authors:

(152) AVP: *no mesmo barriga* 'in the same womb' (Mendes 1985:149)
EP: na mesma barriga

(153) AVP: *cancela pequeno* 'small gate' (*ibid.*)
EP: cancela pequena

(154) AVP: *lixo do nosso cidade* 'the garbage from our city' (Cabral 2005a:48)
EP: lixo da nossa cidade

Nonetheless, the data allow us to conclude that whereas AVP distinguishes between masculine and feminine sex-gender values, there is seldom agreement between the head noun and the non-nuclear elements in the NP, in that there is often a mismatch between the sex-gender of head nouns (i.e. feminine) and that of non-nuclear elements in the NP (i.e. masculine). The same phenomenon also occurs in my corpus:

(155) AVP: *os palavra* 'the words'

(156) AVP: *esses visita* 'those visitors come and go'

(157) AVP: *nesses areas* 'in those areas'

(158) AVP: *todo família* 'the whole family'

(159) AVP: *outro filha* 'the other daughter'
EP: outra filha

(160) AVP: *o mamã* 'mummy'
EP: a mamã

Similar patterns of lack of sex-gender agreement have also been attested in some varieties of BVP, e.g. in that of Helvécia. According to Baxter (in Mello, Baxter, Holm & Megenney 1998:117), "o principal aspecto desta variação tem a ver com a ocorrência de núcleos de SN femininos com modificadores masculinos, por exemplo: *o moço, tia cego*". However, at the present stage of knowledge one cannot affirm that the sex-gender marking rule in AVP's noun phrase is to use masculine modifiers with feminine head nouns, as there are innumerable counter-examples to this rule in the corpus:

(161) AVP: *primeira filho já tem trinta e oito ano* 'the first son is 38 years old already'
EP: o primeiro filho já tem trinta e oito anos.

(162) AVP: *eu tinha **minhas** irmãos* 'my brothers/sisters?'
EP: eu tinha os meus irmãos

(163) AVP: *fui lá em baixo na minha mano, no centro*
'I went to see my brother downtown'
EP: fui lá abaixo à do meu mano, no centro

Hence, rather than lack of sex-gender agreement, what one finds in AVP's noun phrase is the occurrence of variable agreement between the head noun and its determiners. As stated earlier in this section, the existing literature on AVP does not provide any explanations to account for this variation and the interpretation of the data is difficult, as there seem to be no rules governing sex-gender marking and agreement in the NP. In fact, the variation is such that it is often possible to find conflicting strategies of sex-gender marking in the speech of a single informant, which in the case of [+ animate] head nouns such as those in (164) to (167) makes it difficult to evaluate whether the speaker is referring to a male or female entity:

(164) AVP: *agora **a** senhora ficou sozinho! Depois **a** sinhora começa a dizer: "Eu tinha mesmo minhas irmãos e tinha todo **família**.*

(165) AVP: *só neta já tem bué de neta. É mãe do vinte cinco neto!*

(166) AVP: *então, tinha **um avô**, **o meu avô**, avô de pai, que na[s]ceu pai. Po[i]si. Essa **avô** dizia assim (...) não cumpri daquilo **qu'a** velha, **o avô**, dizia*

(167) AVP: *então eu naceu e **o mãe** quando naceu depois me deu o nome ... o meu pai me deu o nome MS.*

This arbitrariness of sex-gender marking and agreement in the NP has also been attested in several other partially restructured varieties of Portuguese, especially in Brazil. For example, Bonvini (2000:402) states that in the ritual "Língua dos Pretos Velhos" not only sex-gender marking but also agreement is random.

The most comprehensive study of similar phenomena of variable sex-gender agreement in the NP in non-fully restructured varieties of Portuguese is Lucchesi's (2000) PhD dissertation on the Helvécia variety of BVP. Building on a variationist approach, Lucchesi (2000) evaluates the ways in which the internal structure of the NP, the morphological properties of elements in the NP and social variables interfere with the variation in sex-gender agreement in the variety's NP. Lucchesi (2000:198-199) concludes that the internal structure of the NP constrains the sex-gender agreement in Helvécia Portuguese in several important ways: (i) simpler NPs (i.e. head noun + a determiner) tend to display agreement; (ii) determiners, modifiers and quantifiers to the right of the head noun tend not to agree with it; (iii) variable head nouns facilitate agreement, whereas invariable ones do not; (iv) the semantic traits of the head noun do not constrain agreement; (v) the existence of number agreement implies the existence of sex-gender agreement. Concerning the morphological properties of the elements in the NP in Helvécia Portuguese, Lucchesi (2000:219) concludes that determiners, rather than quantifiers and modifiers, are the elements that most often receive gender marking, especially those that appear to the left of the head noun or closer to it. With regard to the constraints imposed by social

variables, Lucchesi (2000:236) concludes that younger and middle-aged speakers tend to use sex-gender agreement, as opposed to older speakers; men tend to use agreement more often than women; speakers who spent time outside of the speech community, especially younger speakers, present less variation in agreement; educated or semi-educated speakers use agreement more often than those who are uneducated. Building on these linguistic and social data Lucchesi (2000:219) notes that:

> Na comunidade de fala de Helvécia-Ba, observa-se um uso muito generalizado da regra de concordância de gênero (RCG), o que pode caracterizar uma tendência histórica à fixação dessa regra em termos categóricos na gramática esse dialeto, a exemplo do que já acontece em praticamente todas as demais variedades da língua portuguesa. Desse modo, constata-se na localidade de Helvécia-Ba o desaparecimento iminente de uma das marcas que indicam que essa comunidade de fala passou por um processo de transmissão lingüística irregular durante a sua constituição.

What is interesting about Lucchesi's findings is that - excepting (ii), which holds, and (iii), which is unattested in the present author's corpus - none of them holds in AVP. In fact, the linguistic data show that in AVP (i) both simple and complex NPs, as in (155) to (160) and (152) to (154), respectively, display variable sex-gender agreement; (ii) there does not seem to be any correlation between number and sex-gender agreement, as shown in (162); and (iii) the grammatical category of the non-nuclear elements does not appear to constrain the degree of variation in sex-gender agreement.

The reason why this mismatch is interesting is because it shows that whereas AVP and BVP share important patterns concerning the variation in sex-gender agreement, the greater instability of these features in the former variety attests to its more recent development. Lucchesi's (2000:236) appreciation of the stability of linguistic patterns with regard to BVP is elucidating:

> ... o processo de mudança encontra-se em um dos seus estágios finais. Consideramos que, nesse estágio em que a mudança tende a se completar, opera-se um movimento de homogeneização dos padrões de uso lingüístico dentro da comunidade, diferentemente do que ocorre nos estágios intermediários da mudança, ou nas situações de variação estável, em que são muito nítidas as distinções na forma do encaixamento das variantes na estrutura lingüística e social. Se o momento em que a mudança se completa é aquele em que se uniformizam os usos lingüísticos em toda a estrutura da comunidade de fala, é natural que haja uma tendência a uniformização quando a mudança está prestes a se completar. (*ibid.* 236)

Hence, as far as the NP is concerned, variable sex-gender agreement is possibly the best linguistic evidence of AVP's recent development.

### 4.3.4. Possession

As in EP, possession in AVP is normally indicated by possessive determiners when referring to first or second person possessors. However, when the possessor refers to the third person, the tendency is to use the preposition *de* and the third person marker *ele* (M) or *elas* (F):

(168) AVP: *ele tinha filho* ***deles*** *pequeno.* 'Their sons were small'

(169) AVP: *uma pessoas tinha o irmão* ***dele*** 'One had his brother'

In the formal speech of educated speakers of EP, *seu(s)/sua(s)* is the preferred form to indicate possession, but the use of *de* with the third person marker is also common to avoid any ambiguity regarding the possessor. However, one cannot exclude substrate influence on this structure in AVP. In the Bantu languages surveyed in this chapter, possession is typically indicated by the use of a genitive marker, similar in meaning to the Portuguese preposition *de*, either with the noun or pronoun denoting the possessed thing or person. This is illustrated in (170) with an example from Cokwe (see section 4.2.5 for examples from other Bantu languages):

(170) CC: *kasumbi* ***ka*** *tata* (Martins 1990:50)
chicken POSS POSS;1SG.father
'My father's chicken'
EP: A galinha **do** meu pai

Hence, the use of the preposition *de* in conjunction with the third person markers *ele* and *ela* in AVP are the result of convergent structures in the superstrate and in the substrate languages. However, the word order of possessive determiners in AVP is likely to result mainly from substrate influence. In fact, the structure of the substrate languages demands that possessive determiners (and the genitive particle) always immediately follow the possessed noun regardless of person and number or grammatical categories of the non-nuclear elements in the NP, as illustrated in (171) and (172) below:

(171) CC: *demba we munene* (Martins 1990:164)
rooster your big
'Your big rooster'

(172) CC: *ngombe j'ami jitatu* (*ibid.* 163)
oxen my three
'my three oxen'

The use of possessive determiners after the possessed noun is highly marked in EP in that this is only accepted in very specific syntactic contexts (cf. discussion in section 4.1.3), but in AVP this is the typical word order of possessive determiners regardless of person and the number and grammatical categories of the non-nuclear elements in the NP:

(173) AVP: *eu tenho horário **meu*** 'I have my own schedule'
EP: eu tenho o **meu** horário

(174) AVP: *um dos alunos **seus*** 'one of your students'
EP: um dos **seus** alunos

(175) AVP: *condições que não temos da escola **nossa**.*
'the conditions we don't have in our school'
EP: condições que não temos na **nossa** escola".

(176) AVP: *quioco é a língua **nossa** materna* 'Cokwe is our mother tongue'
EP: quioco é a **nossa** língua materna.

Notice that in (176) the word order matches that of Cokwe in (171) not only with regard to the position of the possessive determiner but also that of the adjective. It is also interesting to notice that possessive determiners typically follow the noun in Portuguese-based creoles, as illustrated below with examples from Cape Verde Creole (CVC) and Anobón Creole in (AC) in the Gulf of Guinea:

(177) CVC: *nɒ ǰa brába di mi*
in island Brava of me
'in my Brava island' (quoted in Holm 1989:274)

(178) AC: *ɔm̃a di'neɲi dɔs ˈlimpi*
hands their two clean
'their two clean hands' (quoted in Holm 1989:284)

The corpus contains other instances which seem to be the result of direct transfer of possessive determiners from Cokwe - cf. (179) and (180) - but since each only occurs once, no definite conclusions can be drawn as to their relevance to an overall description of possession in AVP:

(179) AVP: *N casa é aqui ... a **nosso** casa é aqui* 'Our house is here'
EP: A **nossa** casa é aqui ... a **nossa** casa é aqui.

(180) AVP: *Sim, trabalhar aí **seʒ** casa.* 'Yes, I work there at your place'
EP: Sim, trabalho aí na casa de **vocês**. *or* Sim, trabalho aí na vossa casa. (?)

The origin of *seʒ* in (180) is uncertain, but it is interesting to notice the formal similarity between *N* in (179) and Cokwe's first person singular object person marker *ngu*. In fact, the co-occurrence of Cokwe and Portuguese person markers in general, not only possessive determiners, is worth exploring:

(181) AVP: *Ami **eu** não falo essa língua* [i.e. Cokwe]. 'I don't speak that language'

(182) EP: **Eu** não falo essa língua.

*Ami* is the first person singular marker in Cokwe, which the twelve-year-old boy who uttered this sentences does not speak. However, the first person singular marker in Kimbundu is

formally very similar to *ami* and since the boy was brought up in Luanda this might explain the use of this person marker. The parallel with the use of **ami** as an emphatic particle in some varieties of creole Portuguese is also interesting, as shown in the example from Guinea Bissau Creole (GBC):

(183) GBC: **Ami N** tene un karu (Intumbo 2005)
ENF 1SG 1SG.PRS-ter um carro
EP: Sou eu que tenho um carro *ou* Eu é que tenho um carro.

### 4.3.5. Case and word order of person markers

Clitic direct object forms (*o, a, os, as*) are very rare in AVP. They are systematically replaced either by the corresponding stressed forms (*ele, ela, eles, elas*) as in (184) or by the indirect object clitic form *lhe* as in (185):

(184) AVP: *Deixa **ele** falar!* 'Let him talk!'
EP: Deixa-o falar!

(185) AVP: *É uma sigla porque lemos-**lhe** letra por letra...*
'It's an acronym because we read each letter at a time'
EP: É uma sigla porque **a** lemos letra por letra.

The use of *eles, elas* and *lhe* as direct object person markers is attested in archaic Portuguese (Nunes 1989:238), but whereas the use of *eles, elas* has been maintained in the casual speech of contemporary EP speakers for emphatic purposes (186), especially when preceded by the preposition *a* (187), the use of *lhe* is felt to be ungrammatical even among uneducated speakers of EP [41].

(186) EP: *Olha ele!* (Bechara 2002:175)

(187) EP: *Vi-a **a ela** a passear.*

Still, the presence of these constructions in AVP should not be interpreted as the retention of archaic structures of Portuguese, since the use of Portuguese was not widespread in Angola before the effective colonization of the territory, i.e. in the late 19th century (Gärtner 1989:43) or early 20th century in the case of Lunda Norte and most of interior Angola. It would be better to interpret (184), which is acceptable in EP for emphatic purposes, as having evolved by analogy with the uses in (186) and (187). Considering AVP's tendency to omit word final consonants, (185) seems to be a means of avoiding the occurrence of two contiguous vowels (i.e. *lemo-a*

[41] The use of direct object forms after verb forms ending in r, -s, or –z are problematic for uneducated speakers of EP, who might say "É uma sigla porque lemos-*a* letra por letra...". They would not, however, use the marker *lhe*.

*letra por letra*), which is not possible in most Bantu languages, by using a person marker starting with a consonant, i.e. *lhe*.

AVP also differs from EP concerning the word order of object person markers. In fact, AVP's stressed object forms follow the pattern in EP (i.e. after the verb), but the clitic forms of both indirect object (188) and reflexive and reciprocal (189) person markers clearly diverge from EP word order by typically appearing before the verb in most AVP finite and non-finite sentences:

(188) AVP: *minha mãe e o meu pai* ***me*** *deu o nome de JX ...*
'My parents named me JX'
EP: a minha mãe e o meu pai deram-**me** o nome de JX ...

(189) AVP: *Então, o alfaiate* ***se*** *pendurou ao tronco.*
EP: Então, o alfaiate pendurou-**se** no tronco.

The proclitic use of these person markers was typical of Portuguese until the 17th century, but because the settlement of Angola by the Portuguese was so much later, this does not seem to be a convincing explanation for this feature in AVP. Substrate interference seems more likely, since AVP's word order is similar to the one found in the Bantu languages spoken in Angola: object and reflexive person marker are typically used before the verb. Also, in AVP as in the Bantu languages – cf. (116) - the reflexive or reciprocal pronoun in periphrastic tense constructions is inserted between the auxiliary verb and the root of the main verb:

(190) AVP: *Muitos estavam a* ***se*** *interrogar se a IGM é angolana.*
'Many were asking themselves whether IGM is Angolan.'
EP: Muitos estavam a interrogar-**se** se a IGM é angolana.

The influence of Bantu languages in AVP's use of reflexive or reciprocal person markers is also noticeable in the fact that there seems to be a single form for all persons, i.e. *se*. In fact, with the exception of Kikongo, all the Bantu languages surveyed in this chapter also have a single reflexive marker for all persons and numbers, respectively *–ri-*, *-li-* and *–ri* (cf. table 29 in section 4.2.4).

(191) AVP: *Nós conseguimos se entender.* 'We can understand each other'
EP: **Nós** conseguimos entender-**nos** *or* **Nós** conseguimo-**nos** entender.

Curiously, there are cases in which the reflexive pronoun *se* is omitted, as in (192) and (193):

(192) AVP: *Cansada, sentou_ no caixote* (Mendes 1985:137)
'Tired, she sat down on the box'
EP: Cansada, sentou-**se** no caixote

(193) AVP: Ele chama_ MS. 'Her name is MS'
EP: Ele chama-**se** MS.

Another interesting fact regarding the word order of AVP's object and reflexive/reciprocal person markers is that they tend to appear after the verb in syntactic contexts where EP uses them before the verb (cf. section 4.1.4 for a detailed list of these contexts in EP).

(194) AVP: *a norma é não fazer-se chamada.*
'Usually we do not call out students' names'
(195) EP: a norma é não se fazer a chamada.

However, more extensive data are needed to determine whether this is a matter of perception or a structural feature of AVP.

The reasons for AVP's use of address forms like *você, o senhor / a senhora,* etc with second person markers and verb forms, instead of third person forms as in EP and BVP, is also uncertain:

(196) AVP: *Então, você ficas com este fardo todo em cima de ti.*
'Then, you have to carry this burden'
EP: Então, (**você**) fica com este fardo todo em cima de si.

The most obvious explanation for this is semantic. In fact, like the second person singular marker *tu*, all these forms of address refer to the second person and are therefore combined with second person markers or verb forms (Gärtner 1989:41). However, a more in-depth analysis of the structure of the relevant substrate languages would probably help cast some light on this issue.

## 4.4. Summary and conclusions

The comparison of the structure of the noun phrase in European Portuguese and in Kikongo, Kimbundu, Umbundu and Cokwe shows that Angolan Vernacular Portuguese underwent considerable restructuring of four of the morphosyntactic features that can distinguish partially restructured vernaculars "from the standard variety of their source language" (Holm 2004:140), i.e. number and gender marking as well as agreement, possession, word order and case marking of object person markers.

In fact, unlike European Portuguese, AVP tends to mark number only on the leftmost elements in the noun phrase, which, considering that the head noun typically follows the non-nuclear elements, means that number is seldom marked on it. Hence, as in the Bantu languages, plurality is often expressed in AVP only by means of the agreement markers on the non-nuclear elements in the noun phrase rather than on the head noun itself.

Gender marking and agreement in AVP's noun phrase is even more problematical. In fact, while AVP distinguishes between feminine and masculine, there is often a mismatch between the gender marked on the head noun and that marked on the non-nuclear elements in the NP. Unlike similar phenomena in other partially restructured varieties of Portuguese, (variable) gender agreement in AVP appears to be random.

In AVP, like EP, possession is expressed both by means of possessive determiners and the preposition *de*, but the word order of possessive determiners is closer to that of Bantu languages. The same is true for AVP's object pronouns, which differ from those of European Portuguese in their case marking and word order.

Hence, whereas Angolan Vernacular Portuguese retains many of the morphosyntactic features of its superstrate, it also displays a significant tendency for the kinds of restructuring which tend to make it more like its substrate languages. In chapter 5, these morphosyntactic features will be analyzed within the context of the available sociolinguistic data for language contact in Angola in an attempt to determine how European Portuguese underwent restructuring and developed into AVP.

# Chapter 5

## The development of AVP: conclusions

### 5. 0. Introduction

This thesis is a preliminary study of the vernacular Portuguese that has arisen in Angola, focusing on the social and linguistic factors that shaped its historical development and synchronic structure. This study reviews the different ways AVP has been described in the literature, paying particular attention to the relevance of Holm's (2004) model for varying degrees of restructuring (chapter 2). It goes on to analyze the sociolinguistic setting in Angola from the arrival of the Portuguese to the present (chapter 3), and it evaluates the influence of the relevant languages in Angola in the restructuring of AVP's noun phrase (chapter 4). This concluding chapter attempts to synthesize these finding to ascertain precisely how Portuguese underwent restructuring and developed into a vernacular language in Angola.

### 5.1. The development of AVP: a tentative hypothesis

Chapter 2 reviewed the literature on AVP, showing that there is a growing awareness among linguists that Angolan learners of Portuguese as a second language speak a contact-influenced vernacular whose structure differs significantly from that of European Portuguese. Furthermore, some of the distinctive features of this vernacular have been incorporated into the speech of native speakers of Portuguese in Angola. It is this variety of Portuguese, which is referred to here as Angolan Vernacular Portuguese, that is spreading throughout the country as both a first and second language.

The view that AVP does not differ from EP except in its oral use (essentially at the phonetic and lexical levels, e.g. neologisms and borrowings from Bantu languages) is clearly contradicted both by the early Portuguese grammarians, who indicated that significant language contact phenomena could be found in Angola as early as the 16$^{th}$ century, and by Schuchardt's (1888) evidence of the emergence of a distinctively Angolan variety of Portuguese in the 19$^{th}$ century. This suggests that AVP underwent a significant degree of restructuring due to the influence of

the Bantu languages spoken natively by the majority of the Angolan population. However, most of these studies were based on written texts rather than everyday speech, and they consist largely of enumerations of the lexical and morphosyntactic features that resulted from contact without providing any comprehensive description of the sociolinguistic setting of language contact in Angola.

Concerning the nature of this restructuring, two hypotheses emerge in the literature. One, first suggested by Valkhoff (1966) and more recently supported by Lipski (1995, 2005), is that AVP evolved from the Portuguese-based pidgin used as a lingua franca along coastal Africa from the 15th to the 19th century. This pidgin was the result of both imperfect acquisition of Portuguese by adult speakers of African languages and direct interference from those languages. Due to increased contact with Portuguese, it developed into a language variety closer to the European standard but characterized by sporadic deviations from it. Synchronic features of AVP are then simply the remnants of this old Portuguese-based pidgin.

The second hypothesis, which is supported by the overwhelming majority of the linguists surveyed in chapter 2, suggests that AVP never underwent full restructuring. The features that distinguish it from standard EP are rather the result of the natural drift of earlier Portuguese, the leveling of regional European dialects, and transfer from Angolan languages due to imperfect language acquisition and multilingualism. This started in fifteenth-century coastal Angola and spread to the interior mainly after independence from Portugal, when a largely Bantu-speaking population was forced to acquire some degree of proficiency in Portuguese. According to this view, AVP's synchronic structure is essentially a twentieth-century phenomenon.

According to the second hypothesis, AVP does not seem to fit any of the conventional typologies of language contact outcomes as it is neither a pidgin nor a creole, much less a bilingual mixed language. Rather, it appears to be a borderline case between fully restructured and non-restructured languages, in that it displays a significant portion of the European languages' morphosyntax but also the introduction of substrate and interlanguage features. As shown in section 2.2.3, the theoretical model that best accounts for this type of borderline outcome of language contact is Holm's (2004) partial restructuring.

According to this theory, what distinguishes partially restructured vernaculars from their lexical source languages on one hand and the creoles and post-creoles of the same lexical base on the other is the fact that PRVs emerge in a sociolinguistic setting in which, at the early period of their development, there is a more balanced ratio of native versus non-native speakers of the lexical source language in a community whose populations speak different mother tongues and are in the process of changing to a typologically different target language to which learners have

only partial access due to social restrictions often connected to the sociopolitical dominance of that language's speakers.

After the consideration of these hypotheses in chapter two, chapter three provided an overview of the sociolinguistic setting in Angola from the early days of settlement to 1975. The analysis of the sociolinguistic conditions between the 15th and the 17th century in what is now Angolan territory is not consistent with the hypothesis that a stable pidgin was viable there during that period.

In the Congo, it is unlikely that a pidgin ever developed because the linguistic influence of the few Portuguese who settled there was mainly on the ruling classes. In fact, there Portuguese acted as diplomats rather than traders, as the kings of the Congo did not allow the Portuguese to deal in slaves. Moreover, slave traders were those least likely to need Portuguese as they operated essentially in the Kikongo-speaking provinces of the kingdom or in neighboring countries where closely related Bantu languages were spoken. The general population is also unlikely to have needed to acquire Portuguese: even missionary work, which was their sole contact with Portuguese culture, was the responsibility of the Kikongo-speaking native clergy.

In Angola, the sociolinguistic setting also does not appear to have been favorable to the establishment of a Portuguese-based pidgin. In fact, the Portuguese were unable to carry out their clearly expansionist settlement policy, which was explicit in the charter granting the colony to Paulo Dias de Novais in 1575, as the number of the colony's white settlers, especially females, was always too low. This led to widespread miscegenation and the almost complete dependence of the Portuguese on their bilingual mixed-race children and on Africans, not only to get slaves but also to explore the territory. As the slave trade intensified after the recapture of Luanda from the Dutch in 1648, the dependence of the Portuguese on African and Afro-Portuguese intermediaries increased and the position of Kimbundu as the most widely spoken language in the colony was reinforced.

This sociolinguistic setting was maintained in Angola until the end of the first half of the 18th century, as the number of Kimbundu-speaking Africans and Afro-Portuguese clearly continued to outnumber that of Portuguese speakers. It was only in the 19th century that there was any serious competition to Kimbundu as the most widely spoken language in the colony as a consequence of the significant rise in the number of white Portuguese-speaking settlers that arrived in Angola following the independence of Brazil in 1822 and the Conference of Berlin in 1884. While these attempts at settlement in the 19th century did not produce significant results in the colony as a whole, they did contribute to substantial changes in the demographics in the south and in the power relations of the various ethnic groups in Luanda. In fact, the settlements

in the south presented more balanced ratios of whites to blacks and Afro-Portuguese. In Luanda, the development of an embryonic social life and the creation of some facilities for education and recreation put pressure on the Afro-African elite to acquire proficiency in Portuguese and to use it more.

Nonetheless, Kimbundu was still the most widely spoken language in the colony. In the interior, where there continued to be very little contact with Portuguese, that language is likely to have undergone few changes, but in Luanda it was a Portuguese-influenced variety of Kimbundu that was spoken by the overwhelming majority of the black and Afro-Portuguese population. The linguistic features of this variety, which is referred to in the literature as *mestiço de Luanda,* are consistent with the hypothesis that it resulted from slight structural borrowing from Portuguese.

During the first two decades of the 20th century the Portuguese managed to overcome the resistance of the African states through "pacification wars", and from the 1940s onwards they invested heavily in the establishment of agricultural settlements in the interior. While the Portuguese never constituted more than 5,1% of the total population, they came to constitute as much as 60% of the population in the two most important cities in the south of Angola, i.e. Moçâmedes and Sá da Bandeira. In the cities in the north, such as Luanda, despite their limited education, the Portuguese settlers occupied all the higher offices in public administration and commerce, and slowly enforced the use of Portuguese as the only official language in the colony. Moreover, the Portuguese passed a series of laws, such as the *Estatuto do Indígena*, that forced urban Angolans, both black and Afro-Portuguese, to acquire a greater proficiency in Portuguese. Due to contact between different European regional and social dialects, including the Portuguese spoken by the bilingual Afro-Portuguese elite, this Portuguese is likely to have undergone leveling.

In the outskirts of the cities, the sociolinguistic setting was significantly more complex. On the one hand, the system of forced labor had led thousands of Angolans from different linguistic backgrounds to take refuge there. A considerable number of them are likely to have come from Kimbundu-speaking areas, but many also came from areas where other Bantu languages were spoken. Speakers belonging to the same ethnolinguistic group could continue to use their mother tongue, but unlike previous centuries, it is likely that there was no African lingua franca for speakers from different linguistic backgrounds to use to communicate. Imperfect shift to Portuguese and borrowing from first languages are likely scenarios in this sociolinguistic setting.

In the interior, the majority of the population continued to have little or no contact with Portuguese. The *aldeamentos* where they were forced to live after 1961 and throughout the

1970s probably led to the acquisition of basic proficiency in Portuguese, but on the eve of Angola's independence from Portugal the majority of the population in the interior was still not proficient in Portuguese and only a minority used it with any frequency. Shift to Portuguese in the interior is still an ongoing process, mostly motivated by the fact that after 1975 Portuguese became the country's official language, i.e. the language of education, public administration, religion and the media.

The hypotheses above concerning the sociolinguistic setting in Angola in the surveyed periods are supported by the presence of contact-induced features in the AVP noun phrase discussed in chapter 4. In fact, the comparison of the noun phrase in EP on the one hand, and in Kikongo, Kimbundu, Umbundu and Cokwe, on the other, indicates that AVP underwent considerable restructuring of four of the morphosyntactic features that can distinguish partially restructured vernaculars (PRVs) "from the standard variety of their source language" (Holm 2004:140), i.e. number and gender marking as well as agreement, possession, word order and case marking of object person markers.

Concerning number marking and agreement in the noun phrase, the analysis of the data shows that AVP is closer to the structure of the four Bantu languages examined than it is to that of EP, despite the fact that AVP retains many of the features of the latter (i.e. the singular is formally unmarked as opposed to the plural, which is marked by means of the suffix *–s*). On the one hand, plurality is often marked only on the leftmost elements in the NP, whether they are head nouns or non-nuclear elements. Considering that in AVP, as in EP, head nouns typically follow non-nuclear elements, the head noun is thus seldom marked for number. As shown in section 4.2.2, the frequent lack of number marking on the head noun can also be found in substrate languages, which, in a considerable number of noun classes, display a latent tendency to mark number exclusively by means of agreement markers on the non-nuclear elements. Hence, unlike EP, in both AVP and the substrate languages, number appears to be a property of the NP as a whole rather than of the noun itself.

The hypothesis that the lack of plural marking on the head noun is a survival of a phonological rule of –s deletion in Portuguese loans to the Bantu languages is also worth considering. On the other hand, while number is typically marked only on the leftmost elements in the NP in AVP, the specification of which element is to be marked for number appears to be constrained by the number of elements in a particular NP. In fact, in two-element noun phrases plurality is usually expressed by adding *–s* only to the leftmost element, whereas in three-element noun phrases plurality is typically marked on the first two. Again, there is a parallel in the morphosyntax of substrate languages. In fact, when the head noun belongs to a class which displays either no

distinctive plural or singular marker, number is marked only on the non-nuclear element in two-element noun phrases, whereas in three-element noun phrases both non-nuclear elements are marked for number.

The structure of the noun phrase in AVP also differs from that of EP regarding gender marking and agreement. Whereas AVP, like EP, distinguishes between feminine and masculine nouns, there is seldom agreement between the gender value of the head noun and that of the non-nuclear elements in the NP. While similar phenomena of variable gender agreement have been attested in other vernacular varieties of contact-influenced Portuguese, such as Helvécia Portuguese (HP) in Brazil, the AVP data are not consistent with the explanations heretofore offered to account for them. In fact, gender agreement in AVP appears to be random, as it is possible to find both noun phrases in which the determiners, modifiers and quantifiers are masculine but the head noun is feminine, and ones in which opposite marking occurs. The apparent randomness of gender agreement in AVP is consistent with that found in some other varieties resulting from language shift (eg. Helvécia vs. BVP).

As far as possession in concerned, the present analysis of the data shows that the same strategies are used to mark it in both AVP and EP, i.e. either possessive determiners or the preposition *de* in conjunction with the third-person markers *ele* (M) or *ela* (F). However, possession mechanisms in AVP can differ from those in EP in two important ways. On the one hand, the construction *de* + *ele(s)* appears to be the preferred means of expressing third-person possession in AVP, rather than a means of avoiding ambiguity regarding the possessor as in EP. This is likely to be the result of the convergence of similar structures in both the superstrate and substrate languages, as in the latter the preferred means to express possession and related associations is also by means of a genitive marker, similar in meaning to the Portuguese preposition *de*. On the other hand, possessive determiners in AVP, as in the substrate languages, often immediately follow the possessed noun. Considering that this is a highly marked structure in EP but that similar phenomena are found throughout the Portuguese-based creoles, it is likely that this feature in AVP is the result of direct transfer from the substrate. AVP's use of the Kimbundu and Cokwe first person singular marker *ami* also offers an interesting example of transfer from the substrate languages, but as this only occurs once in the corpus no definite conclusion can be drawn concerning its status in AVP.

Finally, AVP also differs significantly from EP in two important ways. First, AVP displays a significant reduction of person marker case contrasts, as direct object forms are systematically replaced either by the corresponding stressed forms or by the indirect object clitic form *lhe*. This reduction in case contrasts is likely to have resulted from analogy with a similar informal

construction used in EP for emphatic purposes and the need to avoid the occurrence of two contiguous vowels by using a clitic person marker with an initial consonant. Second, unlike EP, AVP's clitic indirect object forms and reflexive and reciprocal person markers appear in preverbal position. In this regard, AVP is closer to the structure of its substrate languages, which typically use these forms immediately before the main verb. Another phenomenon that brings AVP closer to the structure of the substrate languages is the fact that it also appears to have a single form for reflexive and reciprocal person markers.

In short, the present analysis of the structure of the noun phrase in AVP in comparison to the structure of the noun phrase in EP and Kikongo, Kimbundu, Umbundu and Cokwe, shows that while AVP retains many morphosyntactic features of the superstrate, it also displays a significant tendency for the kinds of restructuring which tend to make it more like its substrate languages. Moreover, the analysis of the sociolinguistic setting in Angola during the period of contact suggests that this tendency is not the residue of some stable pidgin from the early days of settlement which due to growing contact with Portuguese developed into a language variety closer to standard Portuguese but characterized by sporadic deviations from it. Rather, it suggests that AVP is the result of language shift reflecting its 20$^{th}$-century social setting.

## 5.2. Further research

Gaining a full understanding of the historical development and synchronic structure of AVP is a long-term enterprise which cannot be accomplished by any single individual. However, future researchers would do well to examine AVP's regional varieties from the perspective of contact linguistics with in-depth analyses of their different sociolinguistic histories. Future lines of investigation might also include not only the study of other contact-induced features in the AVP noun phrase, but also such features in the verb phrase and in clause structure.

# References


ADAM, Lucien. 1883. *Les idiomes Négro-Aryen et Laméo-Aryen. Essai d'hybridologie linguistique*. Paris: Maisonneuve.

AIKHENVALD, Alexandra Y. 2003. *Classifiers: a typology of noun categorization devices*. Oxford: Oxford University Press (Oxford Studies in Typology and Linguistic Theory).

ALEGRIA, Maria Fernanda *et al.* (eds.). 1998-2000. *História da Expansão Portuguesa* (Vol.1): *A formação do Império* 1415-1570. Lisbon: Temas & Debates.

ALEXANDRE, Valentim (ed.). 2000. *O Império Africano: séculos XIX e XX*. Lisbon: Edições Colibri.

ALEXANDRE, Valentim & Jill DIAS (eds.). 1998. *Nova História da Expansão Portuguesa: O Império Africano 1825-1890 (Vol. XI)*. Lisbon: Editorial Estampa.

ALLEYNE, Mervyn. 1980. *Comparative Afro-American*. Ann Arbor, MI: Karoma.

AMARAL, Ilídio. 1962. *Ensaio de um estudo geográfico da rede urbana de Angola*. Lisbon: Junta de Investigações do Ultramar.

1987a. Mbanza Congo, cidade do Congo, ou São Salvador: contribuição para o conhecimento geográfico de uma aglomeração urbana africana ao sul do Equador, nos séculos XVI e XVII. *Garcia de Orta* (Série Geografia), *12* (1-2), 1-40.

1987b. A propósito do recenseamento de Luanda de 1983: novos aspectos demográficos. *Garcia da Orta* (Série Geografia), *12* (1-2), 41-66.

2000. *O Consulado de Paulo Dias de Novais – Angola no último quartel do século XVI e primeiro do século XVII*. Lisbon: Instituto de Investigação Científica Tropical.

ANÇÃ, Maria Helena. 1998. Conhecimentos em português-língua segunda – Cabo Verde e Angola. Carlos Reis (Coord.). *Didáctica da Língua e da Literatura - Actas do V Congresso Internacional de Didáctica da Língua e da Literatura - Volume 1*, 1031-42. Coimbra: Almedina.

ARENDS, Jacques. 1993. Towards a gradualist model of creolization. In Byrne & Holm (eds.), 371-380.

ARENDS, Jacques; Pieter MUYSKEN & Norval SMITH (eds.). 1995. *Pidgins and creoles an introduction*. Amsterdam/Philadelphia: John Benjamins Publications (Creole Language Library, 15).

ASHER, R. E. (ed.). 1994. *The Encyclopedia of Language and Linguistics (Volume 1)*. Oxford, New York, Seoul and Tokyo: Pergamon Press.

1994a. *The Encyclopedia of Language and Linguistics (Volume 3)*. Oxford, New York, Seoul and Tokyo: Pergamon Press.

1994b. *The Encyclopedia of Language and Linguistics (Volume 6)*. Oxford, New York, Seoul and Tokyo: Pergamon Press.

ATKINS, Guy. 1955. A demographic survey of the Kimbundu-Kongo language border in Angola. *Boletim da Sociedade de Geografia de Lisboa, 73-79*, 325-347.

BAL, Willy. 1975. À propos de mots d'origine portugaise en Afrique Noire. In Valkhoff (ed.), 119-32.

1979. *Afro-Romanica studia*. Albufeira: Edições Podeidon.

BARRETO, Manuel Saraiva. 1968. O ensino do português como língua segunda em África. In Carvalho (ed.), 145-57.

BARROS, Agnela. 2002. A Situação do Português em Angola. In Mira Mateus (ed.), 35-44.

BARROS, João de. 1957 [1540]. *Gramática da língua portuguesa*. Lisbon [3rd edition, organized by José Pedro Machado].

BATALHA, Ladislau. 1891. *A língua de Angola*. Lisbon: Comp. Nac. Editora (Propaganda de instrução para portuguezes e brazileiros : bibliotheca do povo e das escolas)

BAXTER, Alan. 1998. Morfossintaxe. In Mello et al., 97-134.

BECHARA, Evanildo. 2002 [1999]. *Moderna Gramática Portuguesa*. Rio de Janeiro: Editorial Lucerna, 37th rev. enlarged ed.

BENDER, Gerald J. 2004. *Angola sob o domínio português: Mito e realidade*. Luanda: Editorial Nzila (Colecção Ensaio – 21) .

BETHENCOURT, Francisco *et al.* (Eds). 1998-2000. *História da Expansão Portuguesa* (Vol. 2): *Do Índico ao Atlântico, 1570-1697*. Lisbon: Temas & Debates.

BICKERTON, Derek. 1981. *Roots of Language*. Ann Arbor: Karoma.

*et al* 1984. The language bioprogram hypothesis. *Behavioral and Brain Sciences* 7, 173-88.

1986. Beyond roots: the five-year test. Paper presented at the Linguistic Society of America Summer Institute, City University of New York.

BIRMINGHAM, David. 1966. *Trade and conflict in Angola – the Mbundu and their neighbours under the influence of the Portuguese 1483-1790*. Oxford: Clarendon Press.

2002. Angola. In P. Chabal (ed.), 137-184.

BOLÉO, Manuel de Paiva & Maria Helena Santos SILVA. 1961. Mapa dos dialectos e falares de Portugal Continental. *Boletim de Filologia*, XX, 85-112.

BOLÉO, Manuel de Paiva. 1955. *Unidade e diversidade da língua portuguesa*. Separata da *Revista da Faculdade de Letras de Lisboa*, tomo XX, 2nd series, n .° 1, 5-28.

BONVINI, Emílio. 1994. Angola: language situation. In Asher (ed.) 1994, 127-28.

2000. "La Langue des 'Pretos Velhos' au Brésil un créole à base portugaise d'origine africaine?". *Bulletin de la Societé de Linguistique de Paris*, t. XCV, fasc. 1, 389-416.

BOXER, C. R. 1963. *Race Relations in the Portuguese colonial empire 1415-1825*. Oxford: Clarendon Press.

1965. *Portuguese society in the Tropics: the municipal councils of Goa, Macao, bahia and Luanda, 1510-1800*. Madison & Milwaukee: University of Wisconsin Press.

1977. *A Mulher na Expansão Ultramarina Ibérica – 1415-1815 – Alguns factos, ideias e personalidades*. Lisbon: Livros Horizonte.

1997. *The Portuguese seaborne empire 1415-1825*. London: Carcanet Press.

BYRNE, F. & John HOLM (ed.). 1993. *Atlantic Meets Pacific*. Amsterdam/Philadelphia: John Benjamins.

CABRAL, Lisender Augusto Vicente. 2005. *Complementos Verbais Preposicionados do Português em Angola* ( Vol. I). University of Lisbon (Masters thesis).

2005a. *Complementos Verbais Preposicionados do Português em Angola – Corpus* ( Vol. II). University of Lisbon (Masters thesis).

CÂMARA JR, Joaquim Mattoso. 1970 [reedited in 1996]. *Estrutura da língua portuguesa*. Petrópolis: Editora Vozes.

CANNECATIM, Fr. Bernardo Maria de. 1804. *Diccionario da lingua Bunda ou Angolense, explicada na Portugueza e Latina*. Lisbon: Impressão Régia.

CARVALHO, José G. Herculano de. 1971. *A difusão da língua portuguesa em África*. Luanda: Universidade de Luanda.

CARVALHO, José G. Herculano de. (ed.). 1968. *I Simpósio Luso-Brasileiro sobre a língua portuguesa contemporânea*. Coimbra: Coimbra Editora.

CASTILHO, A. T. de. 1994. Portuguese. In Asher (ed.) 1994b, 3234-36.

CHABAL, Patrick (ed.). 2002. *A history of post-colonial Lusophone Africa*. Bloomington & Indianapolis: Indiana University Press.

CHATELAIN, Heli. 1964 [1888-89]. *Grammatica Elementar do Kimbundu ou Lingua de Angola*. Ridgewood, New Jersey, U.S.A.: The Gregg Press Incorporated.

CHATELAIN, Heli (ed.). 2001 [1894]. *Folk-tales of Angola – fifty tales, with Ki-mbundu text literal English translation, introduction and notes*. Honululu, Hawaii: University Press of the Pacific.

CHAUDENSON, R. 1979. *Les créoles français*. Evreux: Nathan.

CINQUE, Guglielmo & Richard KAYNE (eds.). 2005. *Handbook of Comparative Syntax*. New York: Oxford University Press.

CINTRA, Luís F. Lindley. 1971. Nova proposta de classificação dos dialectos galego-portugueses. *Boletim de Filologia* (Lisboa), XXII, 81-116.

1983. Nova proposta de classificação dos dialectos galego-portugueses. In *Estudos de dialectologia portuguesa*. Lisbon: Sá da Costa, 117-63.

COELHO, F. A. 1880-86. Os dialectos românicos ou neolatinos na África, Ásia e América. *Boletim da Sociedade de Geografia de Lisboa*. (Republished in J. Morais Barbosa (ed.). 1967. 1-234).

COLE, Desmond T. 1971. The history of African linguistics. In Sebeok (ed.), 1-29.

CORBETT, Greville G. 1994. Agreement. In Asher (ed.), 54-60.

1995. *Gender*. Cambridge: Cambridge University Press.

2000. *Number*. Cambridge: Cambridge University Press.

CORRÊA, Letícia M. Sicuro; Diogo A. de A. ALMEIDA & Renata Sobrino PORTO. 2004. On the representation of Portuguese gender-inflected words in the mental lexicon. *Brain and Language, 90,* 63-73. Retrieved September 25, 2005 from the World Wide Web: www.elsevier.com/locate/b&l

COSTA, António Fernandes da. 1997. *Rupturas Estruturais do Português e Línguas Bantas em Angola*. Para uma Análise Diferencial. Minho: University of Minho (unpublished PhD dissertation).

CRUSE, D. A. 1994. Number and number systems. In Asher (ed.), 2857-2861.

CUESTA, Pilar Vasquez. 1990. O ensino do português enquanto língua segunda em Angola. *Angolê: Artes, Letras, Ideias, 1*, 15-8.

CUESTA, Pilar Vasquez & Maria Albertina Mendes da LUZ. 1971. *Gramática da língua portuguesa*. Lisbon: Edições 70.

CUNHA, Celso e Lindley CINTRA. 1995 [1984]. *Nova Gramática do Português Contemporâneo*. Lisbon: Edições João Sá Da Costa, 11th edition.

CURTO, Diogo Ramada. 1998-2000. A língua e o Império. In Alegria *et al.* (eds.), 414-433.

DEGRAFF, Michel. 2003. Against creole exceptionalism. *Language* 79 (2), 391-410.

2005. Morphology and word order in 'creolization' and beyond. In Guglielmo Cinque & Richard Kayne (eds.), 293-372.

DIAS, Jill. 1991. A administração portuguesa ultramarina entre os séculos XV e XX. In Instituto Português de Arquivos, 15-32.

1998. Angola. In Alexandre & Dias (eds.), 319-556.

DIMMENDAL, Gerrit J. 2002. Morphology. In Heine & Nurse (eds.), 161-193.

ENDRUSCHAT, Annette. 1989. Création lexicale en portugais parlé dans la République populaire d'Angola. In Massa & Perl (eds.), 69-86.

1990. A língua falada como força motriz do desenvolvimento do português angolano. *Angolê: Artes, Letras, Ideias, 1*, 8-11.

1993. Acerca da Colocação dos Pronomes Clíticos no Português de Angolanos e Moçambicanos, sua Problemática no Contexto dos Diferentes e na Aquisição da Linguagem. In *Actas do Congresso da Associação Internacional de Lusitanistas.* Lisboa: LIDEL.

ESTERMANN, Carlos. 1971. Cinquenta contos bantos do sudeste de Angola. Luanda: Instituto de Investigação Científica de Angola.

FAO. 1997. Population density of Angola.

FERNANDES, João & Zavoni NTONDO. 2002. *Angola: povos e línguas.* Luanda: Editorial Nzila.

FONSECA, Fernando Venâncio Peixoto da. 1985. *O português entre as línguas do mundo.* Coimbra: Livraria Almedina.

FREUDENTHAL, Aida Faria. 2001. Angola. In Marques (ed.), 259-467.

GÄRTNER, Eberhard. 1989. Remarques sur la syntaxe du portugais en Angola e au Mozambique. In Massa & Perl (eds.), 29-54.

1997. Coincidências dos fenómenos morfo-sintácticos do substandard do português do Brasil, de Angola e de Moçambique. In Degenhardt, Stolz & Ulferts (eds.), 146-80.

2003. Particularidades Morfo-sintácticas do Português não-padrão do Brasil, de Angola e de Moçambique (Lecture at the Faculty of Arts of the University of Coimbra), March 21.

GILBERT, Glenn (ed.). 2002. *Pidgin and creole linguistics in the twenty-first century.* New York: Peter Lang Publishing.

GONÇALVES, António Custódio. 1985. *Kongo. Le lignage contre l'Etat. Dynamique politique Kongo du XVI$^{ème}$ au XVIII siècle.* Lisbon: Instituto de Investigação Científica Tropical.

GONÇALVES, Perpétua. 1985. *Situação actual da língua portuguesa em Moçambique,* 243-51. Lisboa: Instituto de Cultura e Língua Portuguesa (Separata das Actas do Congresso sobre a situação actual da língua portuguesa no mundo – Vol. 1).

1990. *A construção de uma gramática de português em Moçambique: aspectos da estrutura argumental dos verbos.* PhD Dissertation. Faculty of Arts, University of Lisbon.

2004. Towards a unified vision of classes of language aquisition and change: arguments from the genesis of Mozambican African Portuguese. *Journal of Pidgin and Creole Languages 19:2,* 255-59.

GONÇALVES, Perpétua (ed.). 1998. *Mudanças do português em Moçambique: aquisição e formato de estruturas de subordinação*. Maputo: Universidade Eduardo Mondlane.

GORDON, Raymond G. Jr. (ed.). 2005 [15th ed.]. *Ethnologue: Languages of the World*. Dallas, Texas: Summer Institute of Linguistics International. Online version: http://www.ethnologue.com/

GRIMES, Barbara (ed.). 1996 [13th edition]. *Ethnologue*. Dallas: Summer Institute of Linguistics.

HEINE, Bernd & Derek NURSE (eds.) 2002. *African languages – an introduction*. Cambridge: Cambridge University Press.

HOLM, John. 1988-89. *Pidgins and Creoles*. Cambridge: Cambridge University Press, 2 vols.

2000. *An introduction to pidgins and creoles*, Cambridge: Cambridge University Press.

2004. *Contact Languages: the partial restructuring of vernaculars*, Cambridge: Cambridge University Press.

HYMES, Dell (ed.). 1971. *Pidginization and creolization of languages*. Cambridge: Cambridge University Press.

ILIFFE, John. 1999. *Os Africanos – história dum continente*. Lisbon: Terramar.

INSTITUTO NACIONAL DE LÍNGUAS. 1980. *Histórico sobre a criação dos alfabetos em línguas nacionais*. Luanda: Instituto Nacional do Livro e do Disco.

INSTITUTO PORTUGUÊS DE ARQUIVOS. 1991. *Guia de fontes para a história de África*. Lisbon: Comissão Nacional para as Comemorações dos Descobrimentos Portugues, Fundação Oriente, Imprensa Nacional- Casa da Moeda, 3 Vols.

INTUMBO, Incanha. 2005. Balanta, Guiné-Bissau Creole Portuguese and Portuguese: a comparison of the noun phrase. Paper presented at the Max Planck Institute for Evolutionary Anthropology Meeting held in Leipzig, 3-5 June 2005.

KASACK, Sebastian. 1996. Os musseques de Luanda: problemas do desenvolvimento urbano. *Garcia da Orta, Série de Geografia* 15 (1), 65-78.

KLEIN, Herbert S. 2002. *O comércio atlântico de escravos: quatro séculos de comércio esclavagista*. Lisbon: Editora Replicação (Originally published in English – *The Atlantic slave trade*- in 1999 by the Press Syndicate of the University of Cambridge. Translated into Portuguese by Francisco Agarez).

KI-ZERBO, J. & NIAME, D. T (eds.). 1997. *General history of Africa – Vol. IV: Africa from the twelfth to the sixteenth century*. London: James Currey, University of California Press.

KOUNTA, Albakaye. 1980. Sobre as línguas nacionais. *África - Literatura, Arte e Cultura* 2(10), Ano 2 (Outubro-Dezembro). Lisbon: África Editorial Lda.

LABAN, Michel. 1982. Angola: le témoignage de Uanhenga Xitu. *Bulletin des études portugaises et brésiliennes* 42/43: 45-68.

(ed.) 1991. *Angola: encontro com escritores* (Volume 1). Porto: Fundação Engenheiro António de Almeida.

LEONARD, Yves. 1995. La "Communauté des Pays de Langue Portugaise", ou l'hypothétique lusophonie politique. In *Lusotopie - Transitions libérales en Afrique lusophone*. Paris: ed. Karthala & Lusotopie, 9-16.

LIPSKI, John M. 1995. Portuguese language in Angola: luso-creoles' missing link? Paper presented at the annual meeting of the AATSP, San Diego.

2005. *A History of Afro-Hispanic Language: five centuries, five continents*. Cambridge: Cambridge University Press.

LOJENGA, C. Kutsch. 1994. Bantu languages. In Asher (ed.) 1994, 302-4.

LOPES, Carlos. 2003. Poeiras do caminho. Portugal e o espaço lusófono. *Revista Crítica de Ciências Sociais* 66(Outubro 2003), 165-72. Coimbra: Centro de Estudos Sociais.

LUCCHESI, Dante. 2000. *A variação na concordância de género em uma comunidade de fala afro-brasileira*. Unpublished PhD dissertation. Rio de Janeiro: Universidade Federal do Rio de Janeiro.

MACEDO, Jorge. 2002. Todas as variedades do português são bem vindas na literatura angolana. Interview on the website of the *Liga de Escritores Angolanos* retrieved May 9th, 2004 from the World Wide Web: http://www.uea-angola.org/mostra_entrevistas.cfm?ID=323

MAGALHÃES, Joaquim Romero. 1998-2000a. As incursões no espaço africano. In Bethencourt *et al.* (eds.), 65-81.

MAIA, Pe. António da Silva. 1957. *Licções de Gramática de Quimbundo (Português e Banto) – Dialecto Omumbuim*. Author's Edition.

1994 [1961]. *Dicionário complementar Português – Kimbundu – Kikongo*. Cucujães: Editorial Missões – Cooperação Portuguesa [2nd edition].

MARQUES, A. H. de Oliveira. 1995. *Breve História de Portugal*. Lisbon: Editorial Presença.

1998. Organização geral: Viagens e colonização. In Alexandre & Dias (eds.), 136-165.

2001. Introdução: Aspectos de política geral. In Marques (ed.), 21-30.

MARQUES, A. H. de Oliveira (ed.). 2001. *Nova História da Expansão Portuguesa: O Império Africano 1890-1930* (Vol. XI). Lisbon: Editorial Estampa.

MARQUES, Irene Guerra. 1990 [1983]. Algumas considerações sobre a problemática linguística em Angola. In Instituto de Cultura e Língua Portuguesa, 205-23.

MARTINS, João Vicente. 1990. *Elementos de Gramática Utchokwe*. Lisboa: Instituto de Investigação Científica Tropical.

MARTINS, Manuel Alfredo de Morais. 1958. *Contacto de culturas no Congo português: achegas para o seu estudo.* Lisbon: Centro de Estudos Políticos e Sociais. Junta de Investigações do Ultramar (Estudos de Ciências Políticas e Sociais 11).

MASSA, Jean-Michel & Matthias PERL (eds.). 1989. *La Langue Portugaise en Afrique.* Rennes: Université de Haute Bretagne.

MAW, J. 1994. Bantu languages: syntax and morphology. In Asher (ed.) 1994, 304-11.

McWHORTER, John H. 1998. Identifying the creole prototype: vindicating a typological class. *Language* 74, 788-818.

McWHORTER, John H. (ed.). 2000. *Language change and language contact in pidgins and creoles*. Amsterdam: John Benjamins Publishing Company.

MEEUSSEN, Achiel E. 1967. Bantu grammatical reconstructions. *Africana Linguistica 61 (3)*, 79-121.

MELLO, Heliana; Alan BAXTER; John HOLM & William MEGENNEY. 1998. Capítulo 2: O português vernáculo do Brasil. In Matthias Pearl & Armin Schwegler (eds.), *América negra – Panorámica actual de los estudios lingüísticos sobre variedades hispanas, portuguesas y criollas*, 71-173. Frankfurt-am-Main: Vervuert / Madrid: Iberoamericana.

MENDES, Beatriz Correia. 1985. *Contributo para o Estudo da Língua Portuguesa em Angola.* Lisbon: Instituto de Linguística da Faculdade de Letras de Lisboa.

MENEZES-LEROY, Sílvia. 1989. La structure tri-unitaire dans *Nós, os de Makulusu.* In MASSA & Perl (eds.), 93-106.

MIGUEL, Maria Helena R. P. Santos. 1997. Dinâmica da Pronominalização no Português de Luanda. Luanda: Universidade Agostinho Neto (Undergraduate thesis).

MINGAS, Amélia A. 1998. O português em Angola: Reflexões. In *VIII Encontro da Associação das Universidades de Língua Portuguesa* Vol. 1, 109-26. Macau: Cultural Center of the University of Macao.

2000. *Interferência do Kimbundu no Português Falado em Luanda.* Porto: Campo das Letras.

2002. A importância das línguas angolanas na unidade da família angolana. Retrieved March 28, 2004 from the World Wide Web: http://www.uea-angola.org/artigo.cfm?ID=54

2002a. Ensino da Língua Portuguesa no Contexto Angola. In Mira Mateus (ed.), 45-50.

MIRA MATEUS, Maria Helena (ed.). 2002. *Uma Política de Língua para o Português*. Lisbon: Colibri.

MIRA MATEUS, Maria Helena *et al.* 2003 [5th ed, revised and enlarged]. *Gramática da Língua Portuguesa*, Lisbon: Caminho.

MORAIS-BARBOSA, Jorge. (ed.). 1967. *Estudos Linguísticos Crioulos*. Lisbon: Academia Internacional de Cultura Portuguesa.

1969 [2nd ed. rev.]. *A língua portuguesa no mundo*. Lisbon: Sociedade de Geografia de Lisboa.

MOSELEY, Christopher & R. E. ASHER (eds.). 1994. *Atlas of the World's Languages*. London, New York: Routledge.

MUFWENE, Saliloko. 2000. Creolization is a social, not a structural, process. In Newmann-Holzschuh & Schneider (eds.), 65-83.

MÜHLHÄUSLER, Peter. 1997 [1986]. *Pidgin and creole linguistics – expanded and revised edition*. London: University of Westminster Press (Westminster Creolistics Series – 3).

MÜLLER, Friedrich Max. 1875. *Lectures on the Science of Language*, Vol. I. New York: Scribner, Armstrong.

MUTAKA, Ngessimo N. 2000. *An Introduction to African Linguistics*. Muenchen: Lincom Europa.

NARO, Anthony J. & Maria Marta Pereira SCHERRE. 2000. Variable concord in Portuguese: the situation in Brazil and Portugal. In John McWorther (ed.), 235-55.

NASCIMENTO, J. Pereira do. 1893. *Grammatica do Umbundu ou língua de Benguella*. Author's Edition.

NEUMANN-HOLZSCHUH, Ingrid & Edgar W. SCHNEIDER. 2000. Introduction: degrees of restructuring in creole languages. In Newmann-Holzschuh & Schneider (eds.), 1-18.

NEUMANN-HOLZSCHUH, Ingrid & Edgar W. SCHEIDER (eds.). 2000. *Degrees of restructuring in creole languages*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

NUNES, José Joaquim. 1989. *Compêndio de gramática histórica portuguesa*. Porto: Clássica Editora.

OGOT, B. A. (ed.). 1999. *General history of Africa – Vol. V: Africa from the sixteenth to the eighteenth centuries*. London: Heinemann, UNESCO, University of California Press.

OLIVEIRA, Mário António Fernandes de. 1990. *Reler África*. Coimbra: Centro de Estudos Africanos, Instituto de Antropologia, Universidade de Coimbra.

OLIVEIRA, Fernão de. 1975 [1536]. *A gramática da linguagem portuguesa.* Lisbon: Imprensa Nacional - Casa da Moeda [Introduction and notes by Maria Leonor Carvalhão Buescu].

PADRÃO, F. Cerviño. 1998. *A colonização do sul de Angola 1485-1974.* Lisbon: Author's edition.

PAULO, João Carlos. 2001. Introdução: cultura e ideologia colonial. In Marques (ed.), 30-88.

PÉLISSIER, René. 1997 [2nd ed.]. *História das campanhas de Angola: resistência e revoltas 1845-1941* (2 vols). Lisbon: Editorial Estampa.

PEREIRA, Dulce. 2002. As minorias linguísticas em Portugal. In Mira Mateus (ed.), 147-153.

PERES, João Andrade & Telmo MÓIA. 1995. *Áreas críticas da língua portuguesa.* Lisbon: Editorial Caminho (Colecção Universitária – Série Linguística).

PEREZ, Michel. 2000. Les enjeux de la lusophonie. Retrieved March 18, 2004 from the World Wide Web: http://www.lusotopie.sciencespobordeaux.fr/perez.pdf

PERL, Matthias. 1989a. Algunos resultados de la comparación de fenómenos morfosintácticos del "habla bozal" de la "linguagem dos musseques", del "palenquero" y de lenguas criollas de base portuguesa. In *Estudios sobre Español de América y Lingüística Afro Americana* (Ponencias presentadas en el 45 Congreso Internacional de Americanistas), 369-80. Bogota: Publicaciones del Instituto Caro y Cuervo.

1989b. Le portugais et le créole portugais en Afrique. In Massa & Perl (eds.), 9-27.

PERL, Matthias and Christine HUNDT. 1991. Identidade linguística como contribuição para a identidade cultural: reflexões sobre a linguagem da literatura angolana. *Lusorama* 15 (Junho 1991): 45-53.

REINECKE, John (1937). *Marginal languages: a sociological survey of the creole languages and trade jargons.* Ph.D. dissertation, Yale University. Ann Arbor: University Microfilms International.

RILEY, Carlos. 1998-2000. Ilhas Atlânticas e Costa Africana. In Alegria *et al.* (eds.), 137-162.

RUSSEL-WOOD, A. J. R. 1992. *A world on the move – the Portuguese in Africa, asia and America 1415-1808.* New York: St. Martin's Press.

1998-2000a. Fronteiras de integração. In Alegria *et al.* (eds.), 238-255.

1998.2000b. Os portugueses fora do Império. In Alegria *et al.* (eds.), 256-265.

1998-2000c. Políticas de fixação e integração. In Bethencourt *et al.* (eds.), 126-150.

SAMARIN, William J. 1997. The creolization of pidgin morphophonology. In Spears & Winford (eds.), 175-216.

SANTOS, J. Almeida. 1962. *Gramática comparada dos falares bantos angolanos: Umbundu, Musele, Olunaneka, Kwanama, Kimbundu, Omumbwi, Kikongo, Tchokwe*. Nova Lisboa: s/n.

SANTOS, João Marinho dos. 1998. *Estudos sobre os Descobrimentos e a expansão portuguesa*. Coimbra: Faculdade de Letras da Universidade de Coimbra.

SANTOS, Maria Emília Madeira. 1988 [2nd ed.]. *Viagens de exploração terrestre dos portugueses em África*. Lisbon: Centro de Estudos de História e Cartografia Antiga do Instituto de Investigação Científica Tropical.

SCHUCHARDT, Hugo. 1884. *Slawo-deutsches und Slawo-italienisches*. Graz: Leuschner and Lubensky.

1888. On creole Portuguese. In Schuchardt 1979, 59-72.

1979. *The ethnography of variation: selected writings on pidgins and creoles*, edited and translated by T. L. Markey. Ann Arbor: Karoma.

SEBBA, Mark. 1997. *Contact languages - pidgins and creoles*. London: Macmillan.

SEBEOK, Thomas A. (ed.). 1971. *Linguistics in Sub-Saharan Africa (Current trends in linguistics – Volume 7)*. The Hague, Paris: Mouton.

SIEWIERSKA, Anna. 2004. *Person* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

SILVA NETO, Serafim da Silva. 1992 [1957]. *História da língua portuguesa*. Lisbon: Dinalivro [6th edition].

SPEARS, Arthur K. & Donald WINFORD (Eds). 1997. *The structure and status of pidgins and creoles*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

STROUD, Christopher & Perpétua, GONÇALVES (eds.). 1997. *Panorama do português oral de Maputo* (3 vols.). Maputo: Instituto Nacional do Desenvolvimento da Educação (Cadernos de Pesquisa 22, 24, 27).

TAVARES, José Lourenço. 1915. *Gramática da língua do Congo (Kikongo) – dialecto Kisolongo*. Loanda: Imprensa Nacional de Angola.

TAYLOR, Douglas. 1971. Grammatical and lexical affinities of creoles. In Hymes, 293-6.

1977. *Languages of the West Indies*. Baltimore: John Hopkins University Press.

TEYSSIER, Paul. 1989. *Manual de língua portuguesa (Portugal-Brasil)*. Coimbra: Coimbra Editora.

1997 [1980]. *História da língua Portuguesa*. Lisbon: Livraria Sá da Costa (7th edition).

THOMASON, Sarah Grey. 1997. A typology of contact languages. In Spears & Winford (eds.), 71-88.

1997a. Introduction. In Thomason (ed.), 1-7.

2001. *Language contact: an introduction*. Edinburgh: Edinburgh University Press.

THOMASON, Sarah Grey (ed.). 1997. *Contact languages: a wider perspective*. Amsterdam: John Benjamins Publishing Company (Creole Language Library Vol. 17).

THOMASON, Sarah Grey & Terence KAUFMAN. 1988. *Language contact, creolization and genetic linguistics*. Berkeley, Los Angeles, London: University of California Press.

TIBERIUS, Carole; Greville CORBETT and Julia BARRON. (1999-2002). Agreement: a bibliography. Retrieved October 20, 2005 from *Surrey's Morphology Group* on the World Wide Web: http://www.surrey.ac.uk/LIS/SMG/projects/agreement/agreement_bib.htm

UNITED NATIONS. 1997. Map of Angola (Map No. 3727 Rev.2). New York: Department of Public Information (Cartographic Section). Retrieved March 17, 2005 from the World Wide Web: http://www.un.org/Depts/Cartographic/map/profile/angola.pdf

VALENTE, P. José Francisco. 1964. *Gramática Umbundu – a língua do centro de Angola*. Lisbon: Junta de Investigações do Ultramar.

VALKHOFF, Marius F. 1966. *Studies in Portuguese and Creole: with special reference to South Africa*. Johannesburg: Witwatersrand University Press.

VALKHOFF, Marius (Ed). 1975. *Miscelânea Luso-Africana – colectânea de estudos coligidos*. Lisbon: Junta de Investigações Científicas do Ultramar.

1975a. L'importance do portugais comme langue mondial avant le français. In Valkhoff (ed.), 73-85.

1975b. África do Sul e Portugal. In Valkhoff (ed.), 87-101.

1975c. Presença de católicos e uso do português no Cabo da Boa Esperança em 1685. In Valkhoff (ed.), 103-17.

VANSINA, Jan. 1997. Equatorial Africa and Angola: migrations and the emergence of the first states. In Ki-Zerbo & Niame (eds.), 219-229.

2001. Portuguese vs Kimbundu: language use in the colony of Angola (1575- c. 1845). *Bull. Séanc. Acad. R. Sci. Outre-Mer Mede. Zitt. K. Acad. Overzeese Wet, 47*, 267-81.

VANSINA, Jan & T. OBENGA. 1999. The Kongo kingdom and its neighbours. In Ogot (ed.), 273-289.

VASCONCELLOS, José Leite de. 1987 [1901]. *Esquisse d'une dialectologie portugaise*. Lisboa: Centro de Linguística da Universidade de Lisboa – INIC [3ed ed].

VENÂNCIO, José Carlos. 1996. *A economia de Luanda e Hinterland no século XVIII – um estudo de sociologia histórica*. Lisbon: Editorial Estampa.

VILELA, Mário. 1995. Algumas tendências da língua portuguesa em África. In Mário Vilela. *Ensino e língua portuguesa: léxico, dicionário, gramática,* 45-72. Coimbra: Almedina.

1999. A língua portuguesa em África: tendências e factos. *Africana Studia* 1: 175-95.

VILLALVA, Alina. 2000. *Estruturas morfológicas. Unidades e hierarquias nas palavras do português.* Lisbon: Fundação Calouste Gulbenkian and Ministério da Ciência e Tecnologia.

WALD, Benji. 1994. Sub-Saharan Africa. In Moseley & Asher (eds.), 289-309.

WEINREICH, Uriel. 1953. *Languages in contact: findings and problems*. The Hague: Mouton.

WELMERS, W. E. 1973. *African language structures*. Berkeley: University of California Press.

WHINNOM, Keith. 1971. Linguistic hybridization and the "Special case" of pidgins and creoles. In Hymes, 91-116.

WILLIAMSON, Kay & Roger BLENCH. 2002. Niger-Congo. In Heine & Nurse (eds.), 11-42.

WINFORD, Donald. 1997. Introduction: on the structure and status of pidgins and creoles. In Spears & Winford (eds.), 1-31.

2002. Creoles in the context of contact linguistics. In Gilbert (ed.), 287-354.

2003. *An introduction to Contact Linguistics*. Oxford: Blackwell Publishing.

WORLD BANK. 2004. *Angola at a glance*. Retrieved December 1, 2005 from the World Wide Web: http://devdata.worldbank.org/AAG/ago_aag.pdf