

Universidade do Minho
Escola de Letras, Artes e Ciências Humanas

**A Reflection on the Umbundu Corpus Planning for the Angolan Education System:
Towards the harmonization of the Catholic and the Protestant orthographies**

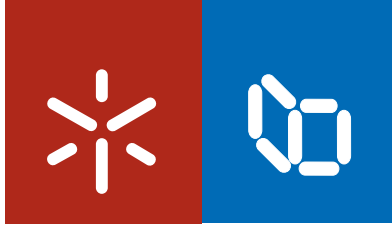
Botelho Isalino Jimbi

UMinho | 2023

Botelho Isalino Jimbi

**A Reflection on the Umbundu Corpus
Planning for the Angolan Education System:
Towards the harmonization of the Catholic
and the Protestant orthographies**

August 2023



Universidade do Minho
Escola de Letras, Artes e Ciências Humanas

Botelho Isalino Jimbi

**A Reflection on the Umbundu Corpus
Planning for the Angolan Education System:
Towards the harmonization of the Catholic
and the Protestant orthographies**

Doctoral Thesis
Doctoral in Language Sciences
Specialization in Sociolinguistics

Work supervised by
Pedro Dono López, PhD
Anabela Alves dos Santos Rato, PhD

August 2023

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.



**Atribuição
CC BY**

<https://creativecommons.org/licenses/by/4.0/>

ACKNOWLEDGEMENTS

This has been a long academic journey that has depended on a plethora of sources of strength to be accomplished.

First, though, I very much thank Professor Pedro Dono López, for his unconditional compliance, guidance, advice, encouragement and, most importantly, the high degree of professionalism and understanding shown, at all levels possible, as a supervisor of this research work. The advices to write articles and indications for publications with credited journals, the supervision meetings both online and presential were incredibly encouraging and very revealing of his peculiar personality. Thanks for being so kind!

I so much feel indebted to Professor Anabela Alves dos Santos Rato who decided to carry on co-supervising this research project with incredible advices on linguistic level (being non-native user of English), despite the fact that she had no longer been institutionally linked to UMINHO, sending feedbacks and support from Toronto, Canada. Words cannot express all the gratitude there is to it.

I also owe a great deal of gratitude to Professor Isabel Cristina da Costa Alves Ermida for having received and positively accepted the research project as a viable one to proceed. I have to acknowledge the professionalism and responsibility in making the transition to Professor Anabela Rato in due time. Your permanent encouragement for me to proceed was a fuel throughout the process of writing.

To Professor Jaime José Becerra Costa I am indebted for having played as key strategist by paving new ways for this research project to come true. There is no wealth that can pay for such magnanimity! You always made it clear that I could do it. Your direction and belief have made me come to the end of this research project today.

I am very thankful to my family, especially my right praiseworthy companion Cláumar Jimbi, for having believed, spurred and supported unconditionally this hard stage of our lives.

The Jimbi, my siblings, I have so much appreciated the unconditional financial and moral support and the willing to make me feel at ease with my studies in the many times I decided quit. I especially refer to Flora Jimbi, Graciana Jimbi, Dina Jimbi, Evaristo Jimbi, Bernardino Jimbi, Tarcísio Jimbi and Rodrigues Jimbi. You have surely fought my fight.

I have been very blessed to have so many precious friends that all cannot be mentioned here: Dinís Vandor Sicala, Vita Kaley, for having advised and given unconditional support, at all levels, at crucial stages of this process. May this spirit of giving in to others' necessities remain in you.

Carlos and Manuela Ribeiro and Dr Helena Vaz, have been of fundamental help in times I needed them much for survival.

Last but not least, I thank my dearest friend Ndjimi Ndumba Watembo Malaka, for having been a companion and a support in the very hardest days.

To all of you, whom I may have forgotten to mention, I express a cordial THANK YOU from the very bottom of my heart.

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

A Reflection on the Umbundu Corpus Planning for the Angolan Education System: Towards the harmonization of the Catholic and the Protestant orthographies

ABSTRACT

In Angola, the government has been trying to implement the teaching of autochthonous languages into the education system. The problem this research work seeks to explain in this context is that Umbundu, a native language of Angola, is written in various main standard orthographies: the Catholic Standard Orthography (hereafter, CSO), the Protestant Standard Orthography (hereafter, PSO), the Jehovah Witnesses Standard Orthography (hereafter, JWSO) and the State Standard Orthography (hereafter, SSO). Two important questions to the problem of this research were considered. The first one is 'What factors have been behind the appearance of various orthographic systems for Umbundu?' The second is 'How to harmonize the orthographic systems of Umbundu so that they can respond to the need to write it in simpler and ameliorated ways resulting from consensus of the users?' To answer the questions above two hypotheses were proposed. The first is that the encoders, having not all been, neither language specialists nor Umbundu native speakers, took advantages of their political and social positions in the colonial epoch and decided to encode the language according to the intellectual capabilities that they had. The second hypothesis is that by exploring the maximum out the various orthographical systems of the world – containing graphs which lack in the Portuguese alphabet – the Umbundu writing system could be made more apt for the needs of the education system. In order to collect the necessary data for the study, a triangulation approach was considered, so that a questionnaire for teachers/trainers and another for members of churches and common users of Umbundu were used. On the other hand, a documents analysis exercise was carried out focusing on the texts from the bibles, on excerpts of the Umbundu coursebook used in State schools and on samples of the translations of the respondents. In general, the research has concluded that, on the one hand, the Umbundu orthography is made of different standards, mainly according to each religious group and, on the other hand, the State has proposed a relatively different orthography without training teachers to use it, making most of these professionals teach Umbundu according to their religious tradition. The general suggestion is to involve language related specialists, language grassroots and governmental agents and get them work out an orthography that represents the harmonization of the Umbundu orthography for the education system.

Keywords: corpus planning, education system, harmonization, language planning, Umbundu orthography.

Uma Reflexão sobre a Planificação de Corpus da Língua Umbundu para o Sistema de Educação de Angola: Uma Tentativa de harmonização das ortografias Católica e Protestante

RESUMO

O governo angolano tem vindo a trabalhar na inserção das línguas locais no sistema de ensino. O problema desta pesquisa, no contexto em que se escreve, prende-se com o Umbundu, uma língua nativa angolana proposta para ser inserida no sistema de ensino, que se encontra escrita e adotada com base em várias ortografias principais: o Padrão Ortográfico Católico (POCA), Padrão Ortográfico Protestante (POPA), o Padrão Ortográfico das Testemunhas de Jeová (POTEJEO) e o Padrão Ortográfico do Estado (POE). Esse problema suscitou duas questões científicas importantes. A primeira delas tenta entender os factores que estão na base do surgimento dos vários sistemas ortográficos para a língua Umbundu. A segunda tenta buscar as formas de harmonizar os sistemas ortográficos existentes do Umbundu para que se possa escrever com mais simplicidade e uniformidade resultante da contribuição e consenso entre os usuários. Para dar respostas às duas perguntas, duas hipóteses foram avançadas respetivamente. A primeira é que, provavelmente, os agentes que codificaram as várias ortografias, não sendo todos especialistas em Ciências da Linguagem nem bons conhecedores da língua Umbundu, aproveitaram-se das vantagens que detinham do ponto de vista político e social no período colonial para decidirem codificar a língua com os meios de que dispunham. A segunda hipótese aponta para a possibilidade de se aproveitar das várias ortografias das línguas do mundo – aquelas que contenham sinais gráficos não usados no alfabeto da Língua Portuguesa (que serve de base principal para se escrever a língua Umbundu) – e fazer da ortografia do Umbundu uma mais simples e apta para o sistema de ensino e aprendizagem. Para a recolha das informações para o estudo foi necessária uma abordagem triangulada, i.e., com o uso de um questionário de professores/formadores e outro designado para os membros das igrejas e falantes não religiosos da língua Umbundu. Esses questionários envolvem perguntas de desenvolvimento, dicotómicas, de escolha múltipla e de escala de Likert. Outrossim, fez-se a análise documental com foco em textos bíblicos das várias denominações religiosas cristãs, em amostras das traduções feitas pelos respondentes e excertos de textos do manual usado pelo Estado para o ensino do Umbundu atualmente em fase experimental. Em suma, por um lado, a ortografia do Umbundu é, hoje, feita, principalmente, de padrões que variam de acordo com os grupos religiosos e, por outro lado, o estado propõe/impõe uma ortografia relativamente diferente sem a necessária formação de professores, fazendo com que cada um desses profissionais ensine o Umbundu de acordo com a tradição ortográfica da sua religião ou aquela que achar conveniente usar. O estudo recomenda que especialistas de ciências da linguagem, guardiões da língua e agentes do estado trabalhem em conjunto para uma ortografia do Umbundu mais simples e harmonizada para o ensino e aprendizagem.

Palavras-chave: planificação de corpus, sistema de educação, harmonização, planificação linguística, ortografia do Umbundu.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iii
STATEMENT OF INTEGRITY	iv
ABSTRACT.....	v
RESUMO	vi
LIST OF ABBREVIATIONS.....	x
LIST OF FIGURES	xi
LIST OF TABLES.....	xiii
LIST OF APPENDICES.....	xiv
LIST OF SAMPLES	xv
DEDICATION	xvi
INTRODUCTION.....	1
1. MULTILINGUALISM IN SOUTHERN AFRICA: NATURE, PROCESSES AND IMPLICATIONS TO AFRICAN LANGUAGES STUDIES	6
1.1. The multilingual nature of the Southern African States	6
1.2. The linguistic heritage from the Portuguese colonialism in Angola	10
1.3. Systematic devaluing of the endogenous languages by the colonial authorities	11
1.4. A brief account of the Angolan government language policy	15
1.5. Umbundu: A language of Guthrie’s zone R	17
1.6. The study of Umbundu in the colonial period	20
2. CORPUS PLANNING WITHIN LANGUAGE POLICY AND PLANNING.....	24
2.1. Language policy and planning: An essential distinction	24
2.1.1. Language policy	24
2.1.2. Language planning	27
2.1.2.1. The four dimensions of language planning	28
2.1.2.2. Agents of language planning	31
2.1.2.3. Factors that determine language planning	31
2.1.2.4. Important principles to consider in language planning	32
2.1.2.5. Corpus planning as a dimension of language planning	33
2.1.2.6. Tasks of codification: The standardization of <i>corpus</i>	34

2.1.2.7. Graphization, grammatication and lexication	37
2.1.2.8. Two crucial criteria guiding graphization	42
2.1.3. Corpus elaboration	43
3. CHAPTER THREE: ORTHOGRAPHY: NATURE, TYPOLOGIES AND PROCESSES	52
3.1. The cenemic and the pleremic orthographies	52
3.2. Shallow versus opaque orthography	53
3.3. First proposed criteria for optimum writing systems: A brief account	55
3.4. Principles guiding orthographical systems in the African context	59
3.5. Creating an orthography description: A practical sketch	64
4. PERCEPTIONS AND USES OF THE WRITTEN UMBUNDU BY RELEVANT AGENTS	73
4.1. Methodological approach	73
4.2. Population, sample and sampling techniques	74
4.3. Research tools	75
4.3.1. Questionnaire on the harmonization of the Umbundu orthography: Design and application	75
4.3.1.1. Group one: Personal background	77
4.3.1.2. Conclusive remarks on personal background	84
4.3.1.3. Group two: Learning the language.	86
4.3.1.4. Conclusive remarks on learning the language	87
4.3.1.5. Group three: The practice of Umbundu 1	88
4.3.1.6. Conclusive remarks on practice of Umbundu 1	92
4.3.1.7. Group four: Practice of Umbundu 2 (translations into Umbundu)	93
4.3.2. Teacher/Trainer questionnaire: Design and application	96
4.3.2.1. Data from the teacher/trainer questionnaire	99
4.3.2.2. Group one: Personal background	99
4.3.2.3. Conclusive remarks on personal background	103
4.3.2.4. Group two: learning the language	104
4.3.2.5. Conclusive remarks on learning the language	106
4.3.2.7. Conclusive remarks on the trainers'/teachers' practice of Umbundu	111
5. A PROPOSAL OF THE HARMONIZATION OF THE HUMBUNDU ORTHOGRAPHY	115

5.1. The harmonization of orthography in subsaharan Africa: A conceptual framework	115
5.2. Chronological development of the Umbundu orthography and key authors	119
5.3. Analytical tables of the evolution of the Umbundu orthography	124
5.4. Practice of Umbundu 3 (translations into Umbundu by respondents)	128
5.5. Conclusive remarks on practice of Umbundu 3 (translations into Umbundu)	129
5.6. Current orthographical approaches to Umbundu	130
5.7. A proposal of harmonization of the Umbundu orthography: Alphabet and word parts	137
5.7.1. Previous attempts to ameliorate the Umbundu alphabet	138
5.7.2. The recent proposal of harmonization of the alphabet for the Umbundu language	138
5.8. The use of diacritics to differentiate homographs	147
5.9. A proposal of the biblical passage of Mathew’s chapter 13, verses 1 to 9	150
6. MAJOR RESEARCH CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCHERS.....	153
6.1. Main conclusions of the study	157
6.2. Suggestions with regard to the population, educational agents and previous research	160
6.3. Suggestions for future work	162
REFERENCES.....	163
APPENDICES.....	182

LIST OF ABBREVIATIONS

BLL	Bantu Languages Linguistics
CA	Current Adoption
CSO	Catholic Standard Orthography
JWSO	Jehovah Witnesses Standard Orthography
NP	New Proposal
PLT	Portuguese Language Teaching
POCA	Padrão Ortográfico Católico (the Portuguese translation of PSO)
POE	Padrão Ortográfico do Estado (the Portuguese translation of SSO)
POPA	Padrão Ortográfico Protestante (the Portuguese translation of PSO)
PSO	Protestant Standard Orthography
SSO	State Standard Orthography

LIST OF FIGURES

Figure 1 : Map of the main languages of Angola.....	16
Figure 2: <i>Umbundu group (R10)</i>	18
Figure 3: <i>Umbundu in its multilingual context of Angola</i>	19
Figure 4: Two maps showing discrepancies in the orthographical representations of local toponymy.....	41
Figure 5: Dictionaries:.....	65
Figure 6: Didactical material for the classroom.....	66
Figure 7: Sounds produced according to characters' and/or diacritics' combinations....	69
Figure 8: Age range among respondents	77
Figure 9: <i>Respondents' place of birth</i>	78
Figure 10: The linguistic repertoire of the respondents	78
Figure 11: <i>Respondents' religion</i>	79
Figure 12: Language used by mothers at home	80
Figure 13: Language used by father at home.....	82
Figure 14: <i>Father's writing skills</i>	83
Figure 15: The family language	84
Figure 16: Experience with African languages at school	86
Figure 17: <i>Preferred orthography</i>	87
Figure 18: Desire to have learnt the Umbundu language	89
Figure 19: The language used in the respondents' church.....	90
Figure 20: Awareness of public schools where Umbundu is taught.....	90
Figure 21: Standard Umbundu Orthography in use at experimental schools	91
Figure 22: Respondents' ideal Standard Orthography of Umbundu.....	92
Figure 23: Variations in the written samples done by the respondents	94
Figure 24: Professional position of the respondents	99
Figure 25: <i>Academic situation of the respondents</i>	100
Figure 26: <i>Respondents' area of specialization</i>	101
Figure 27:: Respondents' religion.....	101
Figure 28: <i>Respondents' place of birth</i>	102
Figure 29: <i>Respondents' age</i>	103
Figure 30: Experience with Umbundu at primary school.....	104

Figure 31: <i>Proficiency and skills in Umbundu</i>	105
Figure 32: How the respondents came to be able to use Umbundu	106
Figure 33: Places where the respondents teach Umbundu	107
Figure 34: Respondents' time of teaching experience	108
Figure 35: The standard orthography each teacher uses	109
Figure 36: <i>Who designs the curriculum?</i>	110
Figure 37: Types of didactical material in use for Umbundu	110
Figure 38: Additional aspects that respondents consider for contribution on Umbundu orthography	111
Figure 39: Xitsonga-Shangani's similarities testifying their relatedness	117
Figure 40: Orthography of the same sounds in the two Kenyan languages	118
Figure 41: Developmental traces of the Umbundu orthography	119
Figure 42: The sounds [ʃ] and [g]	124
Figure 43: <i>The sounds [ɲ] and [ŋ]</i>	125
Figure 44: <i>The sound [tʃ]</i>	126
Figure 45: The sound [dʒ]	126
Figure 46: <i>The sounds [s]; [k]; [w]</i>	127
Figure 47: <i>The sound [j]</i>	127
Figure 48: <i>Variation of the Umbundu orthography in translations according to religion</i>	129
Figure 49: <i>The four ways of the Umbundu orthography per standards</i>	131
Figure 50: <i>Schadeberg's systematic phonetic segments of Umbundu</i>	149

LIST OF TABLES

Table 1: A developmental model for language planning goals.....	29
Table 2: Transposition of the text of Matthew Chapter 13, verse 1 to 9	132
Table 3: Synopsis of how the sound [tʃ] is written according to each standard	133
Table 4: Synopsis of how the conjunctive and disjunctive approaches differ per standard	134
Table 5: Synopsis of how the sound[ŋ] is written per standard	134
Table 6: Synopsis of how nasalized vowels are accentuated.....	135
Table 7: Synopsis of how [dʒ] is written per standard	135
Table 8: The previously accepted alphabets	138
Table 9: The adopted alphabet proposal for the Umbundu language.....	139
Table 10: Demonstration of why a disjunctive orthography would be more functional for the learner	142
Table 11: Demonstration of why a disjunctive orthography would be more functional for the learner 2	143
Table 12: Demonstration of why a disjunctive orthography would be more functional for the learner 3	144
Table 13: Splitting verb + noun agglutinations.....	145
Table 14: Splitting verb + adjective agglutinations	146
Table 15: Splitting verb + preposition/adverbs: towards harmonization (disjunction)..	146
Table 16: Splitting preposition + verb + noun agglutinations	147
Table 17: The use of diacritics to differentiate homographs (disjunctive proposal)	148
Table 18: The use of tilde over letters to differentiate words.....	150

LIST OF APPENDICES

Appendix 1: Questionnaire on the harmonization of the Umbundu orthography	182
Appendix 2: Teacher/trainer questionnaire on the Umbundu orthography harmonization	187
Appendix 3: Written samples from the teacher/trainer questionnaire on the Umbundu orthography harmonization	192
Appendix 4: Written samples from the respondents: standard orthography tendencies	193
Appendix 5: Written samples from the teacher/trainer questionnaire on the Umbundu orthography harmonization (continued).....	194
Appendix 6: Test of Umbundu (CSO) 1	195
Appendix 7: Test of Umbundu (CSO) 2.....	196
Appendix 8: Test of Umbundu (CSO) 3.....	197
Appendix 9: Test of Umbundu (a combination of PSO, JWSO and the SSO) 4.....	198
Appendix 10: Test of Umbundu (a combination of PSO, JWSO and SSO) 5.....	199
Appendix 11: Test of Umbundu (a combination of PSO, JWSO and the SSO) 6.....	200
Appendix 12: Test of Umbundu (a combination of PSO, JWSO and SSO) 7.....	201
Appendix 13: Test of Umbundu (SSO) 8.....	202
Appendix 14: Test of Umbundu (SSO) 9.....	203
Appendix 15: Test of Umbundu (SSO) 10.....	204
Appendix 16: Test of Umbundu (SSO) 11.....	205
Appendix 17: Test of Umbundu (SSO) 12.....	206
Appendix 18: Test of Umbundu (SSO) 13.....	207
Appendix 19: Analytical table of how sounds are coded in the various written works of the Umbundu orthography (author's data)	208
Appendix 20: The three most influential variations of the Umbundu orthography per source	212
Appendix 21: Religion variations of the Umbundu orthography seen from the students' translations (How respondents actually write in Umbundu)	214
Appendix 22: Chronological development of the Umbundu orthography of 4 lexemes + 3 expressions and their proponents	215

Appendix 23: Comparative table of the existing orthographical standards, the ameliorated proposal and the respective remarks on a more disjunctive approach 216

LIST OF SAMPLES

Sample 1: Comments on the choice of standard orthography 192
Sample 2: Problems in teaching Umbundu..... 192
Sample 3: Catholic-like orthographical standard patterns 193
Sample 4: Protestant-like orthographical standard patterns 194

DEDICATION

To:

- The late Jimbi and Nalenga, the ones who gave me Umbundu as the first language, with the deepest gratitude of all.
- The Umbundu community and researchers with due respect and responsibility.

INTRODUCTION

It has been enshrined in the Angolan Constitution that African languages used by the Angolan populations will be protected and supported (Article 19, point 2, p.10, promulgated in 2010)¹. Another important regulation is the Base Law of the Education System (Law 13/01 and Law 17/16)² which establish, under the principle of “Lingua de Ensino” (Language of Instruction), that the State shall promote, secure and create conditions, be they human, technical, material and financial to spread and generalize the inclusion and use of native languages in the curriculum. Moreover, the above legal postulates harmonize with the Universal Declaration of Linguistic Rights (Ouane, 2003: xiii) which advocates the revival, maintenance and use of local languages in the education systems in various countries of the world, mainly the third world. In fact, the Article 9 of the Declaration States that “[a]ll language communities have the right to code, standardize, preserve, develop and promote their linguistic systems, without induced or forced interference” (UNESCO, 1996, p. 7). The introductory part of this research project will comprehend the description of the context of the study, the description of the problem, within the context of its study, the research questions, the hypotheses, the aims and the objectives, the significance and the limitations of the study, the main assumptions of the study and the presentation of the structure of this work.

What constitutes the problem of this study is, on the one hand the fact that Umbundu is orthographically represented by two predominant orthographic systems (i.e. the Catholic Standard Orthography (CSO, hereafter) and the Protestant One (PSO, hereafter) (see Diarra, 2003, p. 342) and, on the other hand, there exists the Jehovah Witnesses Standard Orthography (JWSO) and the State Standard Orthography (SSO, hereafter) which makes it challenging for the government to “impose” it in the education system without a careful, consensual *a priori* orthographic agreement resulting from all the proponents of the above standards.

The consequences of the lack of an agreement among the users of Umbundu are true in the work of many researchers and users today (this issue is discussed with more details in Chapter 4, 5 and 6). For example, Malumbu (2007), Chimbinda (2015) and Kambuta (2021) defend the CSO while IESA (2009) and Daniel (2002) defend the PSO. António, Cuianda & Bonifácio (2012) recognize and propose the SSO which is a tentative mixture of the two standards above, let alone the

¹The Article 19 of the Angolan Constitution of 2010, in its point 2, values and promotes the teaching of the main native Angolan languages together with the most important foreign languages such as English and French.

² The Base Law of the Education system promulgated on December 31st, 2001 and the Base Law of the Education system promulgated on October 7th, 2016.

other alternative defended by the Jehovah Witness Bible (JWSO) and catechism documents writers and other independent researchers (de Oliveira, 2012). These variations of the orthographic systems have serious pedagogical implications. For example, the Catholic Umbundu for “animal” is “*otchinhama*”, whereas the Protestant version is “*ocinyama*” (see Jimbi, 2018). All the above discrepancies are propense to indagations whose answers this research work seeks to answer.

The problem described above indicates that for the education system to be supplied with learner-friendly Umbundu class material, studies must be carried out to understand the reasons behind the appearance of many ways of writing the same language. In this context, two crucial research questions were important in the course of this research project. The first is (1) ‘What sociopolitical, historical, technical and linguistic factors have contributed to the appearance and use of the various orthographic systems for Umbundu?’ The second is (2) ‘How to harmonize apparently contradicting orthographic systems of Umbundu so that they can respond to the need to write it in simpler and ameliorated ways resulting from consensus of the users?’ These questions are supposed to trigger informed answers which can unveil the type of social interactions that existed during the time Umbundu has been codified; reveal the most important facts that happened during the codification of Umbundu and what linguistic background there was that contributed to the current discrepancies of the Umbundu orthography.

As a tentative answer to the first question, it has been hypothesised that (1) the encoders may have not all been neither language specialists nor Umbundu speakers who, for their political and social positions in the colonial epoch, decided to encode the language with the intellectual capacity under their disposal at the time of the colonial occupation. These made them decide on which codification patterns to use no matter whether they had been influenced by previous sources in English, in French or in German about Bantu languages. Given the fact that there most probably were not substantial sources in Portuguese and that the Bantu languages of Angola were all agraphic, i.e. there were not any written cultural *memoire* within the Umbundu peoples’ circumscriptions, most of the principal influences were non-Portuguese literacy. Moreover, technical devices were not developed enough to respond to the demands of the linguistic intricacies of the Umbundu language – a language which was almost totally strange to them.

As to the second hypothesis, it has been forwarded that (2) by drawing on the various orthographic possibilities available for the world’s various writing systems containing graphs which lack in the Portuguese alphabet – which serves as the basis to write the autochthonous languages of Angola – the Umbundu language can be written in a more proficient and effective manner and the

morphosyntactic structures ameliorated for instructional purposes. This approach is believed to facilitate the adoption of newer and simpler characters and rules that may simplify the teaching/learning process of the Umbundu language, process of which will involve a rigorous corpus planning.

This research's aim is to propose a reflection on the Umbundu corpus planning for the Angolan education system, in the short term, and so reach an all-inclusive harmonization, in the long term, of the existing orthographic patterns dispersed according to religion, State-related and independent agents.

As such this study will pursue the following objectives:

- To revise and discuss the various contexts, factors and motivations that made decision-makers embark on language planning and policy and corpus planning, for the education systems mainly in the Southern African context and how they carried out this process at different levels to be successful.
- To describe the attempts made, so far, by the Angolan government towards the integration of African languages (Umbundu included) into the education system.
- To provide a historical background of the procedures underpinning the emergency, development and implementation of the Umbundu *corpus* planning and orthography.
- To describe the current State of the Umbundu orthography, by analysing both existing orthographic systems and identify aspects in which conflicts arise, on the basis of proper research tools (questionnaires, written samples from different sources and other constructed synopsis) of data collection.
- To collect and register the opinions, feelings, perceptions and attitudes of the adherents of the religious groups about the State of the Umbundu orthography.
- To carry out a critical analysis of the various scholars' contributions which may have influenced the way the Umbundu orthography has been constructed, and propose an improved and harmonized orthography of the Umbundu language for the current needs of the Angolan education system.

The importance of this study resides in the fact that African languages, in general, have been relegated to the last stage of priorities in the language policies at the post colonial era. Most countries have adopted the languages of the colonizers: Portuguese in Angola and Mozambique, English in Namibia and Zimbabwe and French in Democratic Republic of Congo, just to adduce a few cases. This condition, for the Angolan situation, has reinforced the stigmatization, in the colonial period, of

the African languages which found themselves banned from use by law in the education system and in official institutions (Gonçalves, 1960, p. 33).

This research work has been designed to have six main chapters. The first chapter discusses the nature, processes and implications of multilingualism in the Southern Africa region. As an initial chapter, it captures aspects related to the reasons why many languages and the people who speak them came to coexist. Moreover, it looks at the various ways different languages were maintained to survive the glottophagic pressure of the colonial languages. Another important aspect of this chapter is the revision of the various stages, players and circumstances that the Umbundu language has underwent in the colonial period. This aspect includes the various attempts that were made to legislate in favour of the maintenance of the autochthonous languages of Angola.

The second chapter deals with corpus planning within the area of language policy and planning. This chapter gives details about the area of language policy and planning, its history, players and processes. The subdivisions of language policy are revised and the most important dimensions of language planning are presented. So, agents, processes, principles, codification, graphization, lexication and related concepts are dealt with in sufficient detail. The importance of this chapter is the provision of important theoretical information which illuminates the understanding of the processes of language planning in the world, in Africa and in Angola.

Chapter three focus on orthography. It revises the various attempts that the orthography has been undertaken in other parts of the world and what processes have been involved. In the course of the chapter, aspects of both shallow and opaque orthographies are explained, as well as the principles underlying orthographical design and the way they can be applied to the Umbundu orthography.

The fourth chapter is about the perceptions, use and usage of the written Umbundu by relevant agents. In this chapter the data is described and various tables showing the percentages and the numbers they represent will be shown. The data has been presented sequentially from more personal data to more professional and practical data in four groups for the questionnaire on harmonization of the Umbundu orthography. The data have been collected and coded for processing and samples of written texts from the Bible were analyzed. Similarly, the teacher/trainer questionnaire has been divided into different groups and structured in the same way the other questionnaire was. Conclusions for each data processing are presented throughout the chapter.

In chapter five, a proposal of the harmonization of the Umbundu orthography is made, including a revision about the historical process of the construction of the Umbundu writing standards, its main players and the possibilities there were to decide the various shapes of the Umbundu orthography. Then, a critical look at the current State of the Umbundu orthography is made and the most influential players and the strength of their research are revealed. Finally, proposals of new orthographical possibilities at the alphabetical and grammatical levels to facilitate learning and teaching are considered.

The sixth chapter presents the main conclusions, the recommendations followed by the list of references. The first part focus on the main conclusions on aspects related to contributions from various scholars in the area of sociolinguistics and related fields. The second part recommends the exploration of the various aspects of this study that have the potential to establish new orthographical paradigms of the Umbundu language which can be simpler, teacher and learner friendly.

1. MULTILINGUALISM IN SOUTHERN AFRICA: NATURE, PROCESSES AND IMPLICATIONS TO AFRICAN LANGUAGES STUDIES

The Southern Africa region is made of multilingual countries. The most important factor for the current multilingualism is the Berlin Conference held among the colonial powers in the period from November 15, 1884 to its closure on February 26, 1885 (Cornelis, 1991; Katzenellenbogen, 1996 & Craven, 2015, just to mention some sources). The way Africa came to be partitioned amongst the imperialists in that conference did not pay respect to the encountered African States' borders. As a result, many different ethnolinguistic groups were submitted to the newly-established geographical configurations as well as to the language of the conquerors.

In the Southern African States, vast territories were mainly occupied by Portugal, England, and Germany (de Blij & Muller, 1997, p. 340). With the occupation came the imposition of the colonialists' languages which had to go along with the various different native languages. While some ethnolinguistic groups (previously found as autonomous "nations") were confined to the same colonized territory, others were submitted to the condition of cross-border linguistic communities who had to share two colonialist powers. For example, in Angola, Umbundu and Kimbundu are languages spoken inside the territory of Angola only (as a Portuguese province "established" by the Berlin Conference), whereas the Oshiwambo language (also Oshikwanyama) is shared by both Angola in the South and Namibia in the North and the Luvale (or Mbunda) is spoken in West Zambia and in East Angola. These are ethnolinguistic groups belonging to the wider *Bantu* group (see Webb & Kembo-Sure, 2000, p. 46). That is the multilingual nature of the countries that constitute the Southern Africa region.

1.1. The multilingual nature of the Southern African States

This section presents an overview of the multilingual nature of countries of the Southern part of Africa: Angola, Botswana, the Kingdom of Eswatini (ex-Swaziland), the Kingdom of Lesotho, Malawi, Mozambique, Namibia, South Africa and Zimbabwe. This brief description of this specific zone is important in that it shows that as a result of colonization all of these now-independent countries have become an agglomeration of various ethnolinguistic groups which have to live and accept each other as one nation under the borders designed by the colonial powers (Webb & Kembo-Sure, 2000, p. 46).

As to Angola, besides Portuguese (spoken by 71% of the population), six more native languages, i.e. Kikongo spoken by 8% of the population; Kimbundu with 8% of the speakers; Umbundu spoken by 23% of the population; Chokwe (often written Cokwe in Fernandes & Ntongo, 2002) with 7% of the population; Mbunda (also called Luvale) with 1% of the speakers and, finally, Oshikwanyama (also written Kwanhama) with 2% of the population) (INE, 2016, p.14). All these languages have been implemented at experimental level – wherever possible – in the public schools since 2004, understood as “effort from the Angolan Institute of Research and Education Development (INIDE)” (Pedro: 2013, p. vii; Jimbi, 2018, p. 901).

Another country of this region with multilingual characteristics is Botswana. In that particular country among the existing 25 mother tongues, Setswana with 90% of speakers is a national language of the country, sharing at the same time the status of official language with English. While Setswana is a language of learning for the primary level only, English is used as language of learning for the three levels of education (Webb & Kembo-Sure, 2000, p. 47; Mokibelo, 2014, p. 422).

There also is the Kingdom of eSwatini (Swaziland) in the region has a population of 1 304 000. The main languages in use in that country are English (with 78,300 speakers) and Swati or siSwati (1,090,000) (Simons and Fennig, 2018; Mordaunt, 1990). SiSwati has been projected for the first 4 years of instruction at public schools in concordance with Language in Education Policy in 2011 (Mkhabela, Nxumalo, & Bhebhe, 2018, p. 129). Apart from these languages, 3 other language are substantially used in Swaziland, i.e. Afrikaans (with 15,600 speakers), Tsonga (with a population of 200, 000) and isiZulu (106,000 speakers) (Simons & Fennig, 2018; Mwakikagile, 2010).

The Kingdom of Lesotho with a population of 2,200, 000 (Eberhard, Simons & Fennig, 2019) is another multilingual territory in the Southern Africa region. The *Ethnologue 2019* presents 5 living languages in use in the Kingdom. English and Southern Sotho are the main languages and they are taught at schools (Bendor-Samuel & Hartell, 1989). According to Dalby (2004, p. 576) and Deprez, Du Plessis and Teck (2001: 175), Sotho is the national language of Lesotho, and it was ratified as such on the on the 12 September 1966 by the National Assembly of Lesotho (Rosenberg, Weisfelder & Frisbie-Fulton, 2005, p. 319). As to the other three (minority) languages of the Kingdom, Zulu is used by 248, 000 people, Phuthi is spoken by 43, 000 people and 18, 000 people speak Xhosa (Lewis, 2009).

Malawi – another country of the region – is a big country with 18, 622,000 people (UNDESA, 2017, p. 19). There are 17 living languages spoken in Malawi. 13 of these languages are indigenous, 4 are non-indigenous (Eberhard, Simons & Fennig, 2019). English and Chichewa are co-official

languages with the latter (9,690,520 users) being widely spread as national language and the only endogenous language taught at schools. Chitumbuka (1,546,000 users), Chiyao (1,546,000 users), Chilomwe (2,290,000 users) are the local languages with a significant number of speakers (Nishioka & Durrani, 2019; Kretzer & Kumwenda, 2016). There are other languages, such as Chitonga, Chingoni and Chisena, whose number of speakers represent less than 4 % of the population (Matiki, 2002, p. 1).

Mozambique is the most oriental Southern African country. Its population is estimated of 28,011,000 with 43 listed living languages. Most of the languages are indigenous (41), 3 of which are used in institutions, 6 of them are endangered 11 are vigorously in use and 2 are not originally of Mozambique (Eberhard, Simons & Fennig, 2019). 16 indigenous languages, and counting, (plus Portuguese which has been adopted as official language since independence in 1975) have been selected for language of instruction up to grade 7 since 2002 (Chimbutane & Benson, 2012). The indigenous language with the highest number of speakers is Makhuwa with 3,220,000, followed by Tsonga with 2,780,000 users, and by Lomwe with 1,660,000. Portuguese comes at the fourth position with 1,580,000 users. The fifth place is taken by Sena with 1,390,000 speakers, followed by Chopi spoken by 760,000 people and Tswa with 695 speakers. The eighth most spoken language is Chuwabu, spoken by 664,000 people followed by Nyanja spoken by 599,000 Mozambicans. The tenth indigenous language to consider here is Ndaou, spoken by 500,000 people, and then Ronga which comes with 423,000 native speakers, Makonde with 360,000 users, Nyungwe with 262,000, Tonga with 224,000 speakers, Yao with 195,000 speakers and the last language of instruction to consider here is Mwani spoken by 100,4000 people.

Namibia is a country situated at the Southwest side of Africa, with a population estimate of 1.6 million (Webb & Kembo-Sure, 2000, p. 49; Legère, Trewby & van Graan, 2000, p. 1) and updated as 2,533,224 by July 2018 (as in the 2019 CIA World Factbook). According to the previous source, Oshiwambo languages (Oshikwanyama, Ngandyera, Mbalanhu, Kwambi and Ndonga) (Legère, Trewby and van Graan, 2000: 26) are spoken by 49.7% of the population, Nama/Damara spoken by 11%, Kavango languages (Rukwangali and Rugciricu/Rumanyo) used by 10.4% of the Namibians, followed by Afrikaans with 9.4% of the people. Herero languages (Otjiherero or Oshierero and Zemba) occupy 9.2% of the Namibians, Zambezi languages, including Cisubiya, Chifwe, Shiyeyi, Chitotela, Mbalangwe and Silozi (Steigertahl, 2018, p. 137) occupy 4.9% of the Namibians, English (the official language) is represented by 2.3% of the population, other African languages represent 1.5%, and other European languages represent 0.7% (see the 2019 CIA World Factbook).

Not all these languages are used as media of instruction in Namibia. In fact, Legère, Trewby and van Graan (2000, p. 27) State that

for large contingents of learners' medium of instruction and learning is a linguistic variety which is often rather closely related, or even mutually intelligible, with the variety spoken at home, for example Oshindonga which has been imposed on all learners outside Ondonga and Oukwanyama, or Silozi in Caprivi which is used as medium of instruction among Sifwe, Siyeyi and Cisubiya speakers.

The above quote clarifies that the inclusion of all languages constitutes a very difficult chore to tackle. However, there has been a demonstrated effort to, at least try to include those languages which can serve as *lingua franca* for those languages which are closer in terms of intelligibility with each other. This State of affairs may also indicate that the availability of resources has its weight when it comes to choose which language should be included and which one should not.

South Africa is one of the most populated countries in the Southern Africa. The *Ethnologue 2019* attributes South Africa 56, 000, 000 people as of 2016 data (Eberhard, Simons & Fennig, 2019). 11 principal languages, out of the 34 living languages, have been selected as official languages (Ngcobo, 2007, p. 2). According to the *2019 CIA World Factbook*, in terms of percentages allocated to the official languages, the most spoken is isiZulu with 24.7%, followed by isiXhosa with 15.6%, Afrikaans with 12.1%, Sepedi with 9.8%, Setswana with 8.9%, English with 8.4%, Sesotho with 8%, Xitsonga with 4%, siSwati with 2.6%, Tshivenda with 2.5%, isiNdebele with 1.6%. The other non-official languages (including Kofi, Nama, and San languages) correspond to 1.9% of the population (Eberhard, Simons & Fennig, 2019). In fact, it has been claimed that the South African language policy and planning is one of the best in the world, although relevant researchers still question the level and the dynamics of its implementation (Ngcobo, 2007).

Zimbabwe is the last country of the Southern Africa region whose multilingualism is under the same description as the others. It is a country of 14,030,368 people, according to estimates of July 2018 (2019 CIA World Factbook). The number of living languages, in Zimbabwe, is of 22 (Eberhard, Simons & Fennig, 2019). The Shona language is the most broadly spoken language and it has acquired a status of official language with a well-established orthography (Hungwe, 2007, pp. 139 – 140). The second most spoken language is Ndebele, which is also an official language with a well-developed orthography and used as language of instruction (Hungwe, 2007, p. 139). English is a principal language used for foreign affairs and official business. As it can be inferred, English is not an indigenous language of Zimbabwe. There are other 13 languages which have also been given the status of official languages. They are: “Chewa, Chibarwe, Kalanga, Koisan, Nambya, Ndau,

Shangani, sign language, Sotho, Tonga, Tswana, Venda, and Xhosa” (2019 CIA World Factbook; Bendor-Samuel & Hartell, 1989; Campbell & King, 2011). It is important to state here that of the 22 living languages 6 are endangered and 1 is dying already (Eberhard, Simons & Fennig, 2019).

The above brief revisiting of the multilingual nature of the countries of southern region of Africa helps understand the state of language attitude, choice, policy and planning within those contexts. Moreover, it serves as reference base to understand the why and the how of some countries to having been more assertive than the others in terms of including languages into the education systems after their independences from the colonial control and the hardship and complexities the process engenders. In details, the next section tries to look at the background information that there is to explain and understand the meanders which paved the state of the Angola language policy and planning in the colonial period.

1.2. The linguistic heritage from the Portuguese colonialism in Angola

As presented above, there have been ethnolinguistic communities spread over what is today the Republic of Angola. In fact, research estimates that there have been inhabitants, in this territory, for at least 12000 years (Fernandes & Ntongo, 2002, p. 17). In 1483 the first contact between the Portuguese and Angolans (at the coastal province of Soyo) took place and by 1491 King Nzinga a Nkuwu and all the chiefs around him had been converted and baptized Catholic (Heywood, 2017, pp. 14 – 15). With the Catholicism came all the influence from the European culture. Research reveals that between 1604 and 1608 Europeans who met the Congolese ambassador in his voyages to Portugal, Spain and the Vatican got surprised by his proficiency in both Portuguese and Latin (Heywood, 2017, p. 15).

The legacy from the Portuguese colonialism has been as immense as expected from any dominant power over a territory. Apart from the cities, railways and the Christian faith, one of the legacies that the Portuguese colonialism has left for Angola is the “language” (Heywood, 2017, p. 15). Portuguese as a language has made it possible for the Angolan natives to code their revolutionary actions in a more elaborated way towards independence. However, the Portuguese authorities, in the era of colonization, did not see the use of native language and culture as a means through which a peaceful dialogue could take place (Gonçalves, 1960, p. 33). In fact, at the very beginning of the occupation, Queen Njinga Mbandi of the Matamba kingdom (when she still was a princess) saw her name being changed into Anna de Sousa, a European one (Pinto, 2014: 11). As witnessed in Rocha (2011, pp. 856- 857)

...do batismo, sinaliza, no context [sic.] socioeconômico, cultural e colonial de Angola, na metade do século XVII, um processo de transformação simbólica e real, em que Jinga “renega sua crença” nas divindades de seu povo, e assume a religião monoteísta do invasor. A princesa mbundu, que nunca aprendeu o idioma português, recebeu, após a cerimônia, uma identidade lusa: o nome Anna de Sousa, Anna em homenagem à madrinha Anna da Silva, e Sousa, como uma espécie de filha adotiva do governador. O rito do batismo cristaliza a entrada da bárbara antropófaga no mundo civilizado cristão europeu.³

As reported above, native people had to be civilized and see their names and identities changed to conform to the desire of the colonizer. This shows the real purpose of the colonization of Angola which included the eradication of all forms of native cultural manifestations, including the value of native languages shaped in their original names. One consequence of neglecting the use of native names is the impediment of its written form for the posterity by means of a proper developed orthographical system.

The following subsection devotes some lines on the various different ways languages were submitted to a derogatory position during the colonization of Angola. It will be seen that the most important decisions were made by the force of the colonial law, and that indigenous languages were devalued to the limit.

1.3. Systematic devaluing of the endogenous languages by the colonial authorities

The devaluation of indigenous languages by the Portuguese colonizer had been recognized by José Pereira do Nascimento, a medical doctor in his *Grammatica do Umbundu or Lingua de Benguella*, where he compares the colonization approaches of Portugal to the ones used by England and France and Germany with regard to the recognition of the importance of the native languages (Jimbi, 2018, p. 901 – 902). The medical doctor above lamented that,

Pena é que Portugal...não possua grammaticas e dictionarios sobre as mais importantes linguas das suas colónias. Comparados com os numerosos trabalhos d'esta ordem realizados modernamente em França, Inglaterra e Allemanha... (Do Nascimento, 1894, p. X).⁴

³ The quote translates as follows:

...from the baptism, he States, in the Angolan socioeconomical, cultural and colonial context, in the half of the 17th century, a process of symbolic and real transformation in which Jinga (the same as Nzinga), disowns her faith in the divinities of her people, and assumes the monotheist religion of the invaders. The *Mbundu* princess, who had never learnt the Portuguese idiom, received, after the ceremony, a Lusitanian identity: the name Anna de Sousa, Anna in reverence to the godmother Anna da Silva, and Sousa as a kind of adoptive daughter of the governor. The baptism ritual crystallizes the initiation of anthropophagic barbarian into the European Christian civilized world.

⁴ The above quotation translates as follows: It is pity that Portugal...does not have grammars and dictionaries about the most important languages of its colonies in comparative terms with the ones carried out by France, England and Germany.

By 1894, Portugal had occupied Angola for more than 450 years without any demonstration of interest in the use of native languages so as to facilitate the colonized to have their culture registered in the language they understand and through which they could be better understood as humans with a history to be coded in print. Do Nascimento exposes the relative despise of the Portuguese colonizer in terms of the establishment of a peaceful relationship with the people whose potential to contribute with knowledge to the world was of great advantage to take.

The intention of making the peoples of Angola not allowed to use their own language was of Norton de Matos who imposed a unique language, the language of the colonizer, for all (Silva, 2008, p. 368). For this intent he had to strategically involve both the Protestant and the Catholic churches as observed by Professor João Pereira Neto,

Mas, para esse efeito era necessário emendar aspectos da organização missionária cristã – Protestante, de modo a garantir que o ensino fosse exclusivamente em língua portuguesa. Neste caso concreto, entre outras vantagens, foi concedido um subsídio anual de três mil escudos a cada missão que tivesse em serviço permanente um professor europeu, missionário ou não, que possuísse as condições e habilitações necessárias para o bem ensinar a língua portuguesa. Uma vez que esta vantagem era atribuída não só às missões Protestantes, mas também às católicas, Norton de Matos de forma hábil e discreta, assegurou o ensino da Língua Portuguesa em Angola (Neto, 2011, p. 100).⁵

It can be seen that the then High Commissioner Norton de Matos was the first to establish a clear and well budgeted language policy which would come to influence language attitudes towards not only Portuguese but also the native Angolan languages. The famous Decree 77, published in the Angola's Official Bulletin (*Boletim Oficial de Angola, 1921*), Nr 5, 1st series (December 9, 1921) Stated that native languages (and foreign languages) had to be banned from classrooms whatsoever. To understand the force of this law one would want to read extracts of it as follows, with our translation into English (also see Gonçalves, 1960, p. 33; Zau, 2010, p. 406⁶; da Costa, 2015, p. 22):

Point 3: The teaching of Portuguese shall be mandatory in all missions;

Point 4: The teaching of any foreign language shall be forbidden;

⁵ The translation of the above quotation goes that

...to this end, amendments of Christian missionary aspects – Protestants, in a way that teaching could only be conducted in Portuguese had to be taken into account. In this concrete case, among other advantages, an annual subsidy of three thousand Portuguese coins was granted for each mission having in its permanent service a European teacher, missionary or not, so long as they could demonstrate to have conditions and education necessary for qualified teaching of the Portuguese language. Once this advantages had been given not only to the Protestants but also to Catholic missions, Norton de Matos skillful and carefully reassured the teaching of the Portuguese language in Angola.

⁶ See also page 6 at <http://unia.ao/docs/FilipeZau-AssisJunior.pdf>

Art. 2: The teaching of indigenous languages shall not be allowed in the missions' schools;

Art. 3: The use of indigenous languages shall only be allowed in spoken form in catechesis and, as auxiliary, in the period of the elementary teaching of the Portuguese Language.

Paragraph 1: In the missions' catechesis, in schools and in any kind of relationships with indigenous people the use of indigenous languages or others than Portuguese shall be forbidden, be they in written or in spoken forms, in form of leaflets, newspapers, independent sheets and manuscripts whatsoever;

Paragraph 2: Religious teaching books shall not be allowed in languages others than Portuguese, with the only allowance of being accompanied by the text in a parallel version in indigenous language.

Paragraph 3: The use of the spoken language that this article refers to, as well as that of the version of the indigenous language, in the terms of the aforementioned paragraph, shall only be allowed on a transitional basis, and while the knowledge of the Portuguese language remains not widespread among the indigenous, depending on the missionaries the successive substitution of the indigenous languages by the Portuguese language, and as much as possible, in all relationships with the indigenous in the catechesis.

Art 4: The provisions of the two previous articles do not impede linguistic endeavour or any other scientific research, and the government reserves its right to prohibit their circulations when, by means of an administrative enquiry, it can be found out that the public order and citizens and indigenous' freedom and security may be disturbed.

As it can be seen, the subjugation of the peoples of Angola during the colonial period affected the State of the use and preservation of the indigenous languages in a negative way, as it was prohibited by law that these idioms should not be used in education. This may have contributed much to the loss of pride of the endogenous languages, as the most important language for living as a civilized person (or an *assimilado*, at least) was Portuguese, the language of the colonial power (Noré & Adão, 2003: 104). That decree fell so hard onto the Protestant missions whose teachers were, in grand number, non-Portuguese speakers that on the 15th and 16th of May 1922, the

Evangelical (i.e. Protestants) missions conglomerated at Ndondi, in the municipality of Cachiungo (ex-Bela Vista) in Huambo province, to discuss, among other issues, the possibility of the High Commissioner, Norton de Matos, to “reconsider the question of the use of the vernacular in school work”, at least for “the teaching of reading, writing and study of the scriptures” which was not taken into consideration (Neto, 2012, pp. 143 – 144; see also de Oliveira, 2015, p. 71).

As autochthonous languages had been prohibited by the Portuguese authorities of the epoch, their survival would only depend on the effort of missionaries who soon learnt that the best way to transmit the Christian message to Africans was to consider the study and use of native languages. However, the missions were various and the approaches to code the languages depended on the way each colonizer coded their main languages of communication (Schroeder, 2008). This situation resulted in a random way of orthographical representation of the African languages.

For the case of the Umbundu language, there emerged 4 main ways of writing systems: the PSO by the Protestant Church, the CSO by the Catholic Church, the JWSO by the Jehovah Witnesses Church and the nowadays experimental SSO by the State. Although the State has decide to include the study of the African languages of Angola into the education system, there still are other valuable alternatives on the part of stakeholders such as relevant church members and independent researchers, depending on whether they are Protestant members, Catholic members, Jehovah Witnesses members or whether they belong to the State’s orthographical initiatives (See Kambuta, 2021, de Oliveira, 2012). These other orthographical proposals show incoherencies with the potential to complicate the existence of a uniform learning and teaching process in the education system.

Taking all the above into account, language planning at the level of the corpus is found important not only to alert for all-inclusive approach to the Umbundu language planning at the level of the *corpus*, but also to reach an all-considered harmonization of its orthograph to make it apt for the education system. Naturally, the existence of many orthographical approaches may be a sign that there has not been, as yet, an all-inclusive agreement of the users about the existing points of strangulation of the orthography of the Umbundu language.

All the constrains observed above make this study very important in that the insertion of native languages into the education system implies the existence of a harmonized orthography that can respond to the pedagogical needs in the language classroom. For this harmonization to happen, many stakeholders have to be taken into account. Teachers, historians, sociolinguists and the like

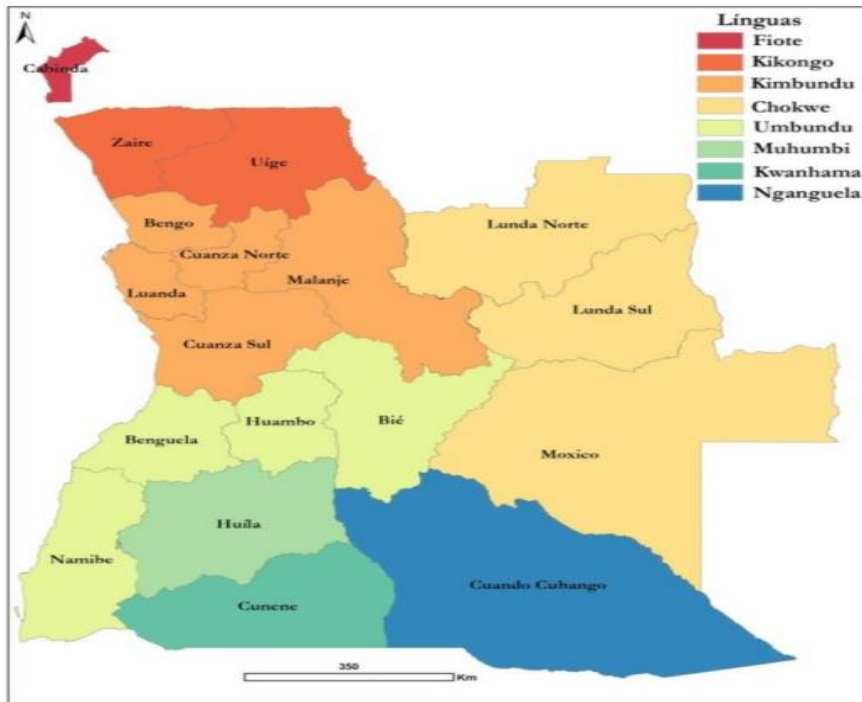
should contribute to the construction of an orthography that can be accepted by all the members of the Umbundu community of speakers for its representational merit.

All in all, the consequences of the language policies in the colonial time had considerable negative effects on the development of Angolan native languages. For the case of Umbundu, in particular, they affected the elaboration of school materials, mainly with the famous Decree 77 by Norton de Matos who determined the use of Portuguese as the sole language of instruction.

1.4. A brief account of the Angolan government language policy

Angola is made of many language communities. After the independence, the government decided to embrace an exoglossic language policy as duly discussed below (cf. Lodhi, 1993; Ruiz, 1995; Zau, 2005 and Zau, 2011). This governmental option made Portuguese the only language of instruction and governmental operations for this long (Ferguson & Heath 1981, p. 531), making the autochthonous languages be relegated to a more or less subordinated status. So, Portuguese stood out to be the language used in the court, in political rallies, the economy, international relations, in the army and in the education system (cf. Eastman, 2001; Mateus & Villalva, 2007). The motivations for the Angolan government to adopt Portuguese were essentially the fears that to consider the insertion of many, not well studied languages, into the official affairs would cause the division of the country into different tribal groups and hinder national unity, schooling, territorial sovereignty defense, the creation of a common front against the invaders and develop an Angolan identity that only Portuguese as a new nation's language could help do (Andrade, 1962; Jorge, 1998; Fry, 2001; Luansi, 2003; Firmino, 2005; Inverno, 2011; Batsikama, 2015; Gaio, 2016; Kajibanga, 2000).

Figure 1 : Map of the main languages of Angola



Note. A map of the main languages of Angola as presented in the report of the census carried out in 2014 (INE, 2016, p. 51).

Figure 1 show a multilingual map of Angola considered in the last Census carried out in 2014. As it can be seen, the map has different colors for each of the 8 languages, namely Fiote (spoken in the province of Cabinda), Kikongo (spoken in the provinces of Zaire and Uige), Kimbundu (used by people in the provinces of Bengo, Cuanza Norte, Malanje, Luanda and Cuanza Sul)), Cokwe (in the regions corresponding to Lunda Norte, Lunda Sul and Moxico provinces) Umbundu Used in the provinces of Bié, Huambo, Benguela and Namibe), Muhumbi spoken in Huila province), Kwanhama (spoken in Cunene province) and Nganguela (in the province of Cuando Cubango).

The ideological background context of the support of the Angolan languages, other than the Portuguese language, can be attributed to Doctor Agostinho Neto, the first president of the independent Angola, often mentioned to have postulated in 1977 that

[T]he exclusive use of Portuguese as an official language, used in our literature, cannot solve our problem...be it in the primary education, or probably in the secondary education, our languages should be used. And given their diversity in the country, sooner or later we must be inclined to agglutinate some dialects to facilitate the contacts...In these passages, the first president showed the desire to embark on the study of our own languages [Angola's autochthonous languages, *emphasis mine*] so that they could be used and implemented in the system of education as languages of instruction (Jimbi, 2018, p. 477ff).

In order to conform to what constituted the desire of the president with regard to the development of the local languages, the National Institute of Languages (ILN, hereafter) came to existence in 1978 under the “Decree No. 62/78 of April the 6th”⁷ This Institute was made to coordinate language-related matters, the research and teaching of the Portuguese Language, the autochthonous languages and the foreign languages teaching and teacher training. However, as time went by the NIL came to its end, and a new denomination was created – the National Languages Institute (NLI, hereafter) – under the Decree No. 40/85 of November the 18th”⁸ The NLI was created to deal with the teaching, promotion, valuing and diffusion of the Angolan native languages, towards their use in various domains.

In the fulfillment of its competences deliberated under the Resolution of the Council of Ministers No. 3/87 of the 23rd of May, the NLI approved the alphabet and transcription rules of six selected autochthonous languages so that they could be tested for the education system in Angola. The *a priori* selection of languages included *Cokwe*, *Kikongo*, *Kimbundu*, *Kwanhama*, *Mbunda* and *Umbundu* (Jimbi, 2018, p. 901). Interestingly, 3 of the languages found in Figure 1 are not included in the list of languages for the experimental stage in the education system (i.e. Fiote, Muhumbi and Nganguela). On the other hand, since 2004, the Ministry of Education has made additional effort to implement the native languages in schools, without considering the essential steps towards embarking onto a language planning process (see Holmes, 2008; Wiley, 1996). This effort from the Angolan Institute of Research and Education Development (INIDE, in Portuguese) has aimed to introduce the teaching of Umbundu, Oxikwanyama (Kwanhama), Nhaneka, “Ngangela”, Kimbundu, Kikongo and Cokwe (Ponso 2011, p. 9). Of the aforementioned languages, Umbundu is spoken by about 23% of the population of Angola (INE, 2016, p. 51).

1.5. Umbundu: A language of Guthrie’s zone R

Umbundu is an Angolan Bantu language classified as belonging to the Niger-Congo phylum (Childs, 2003, p. 219). In fact, Guthrie’s (1948) classification places Umbundu in the R10 zone⁹, and Schadeberg (1990) places it in the R11. About 23% of the Angolan population speaks it (INE, 2016,

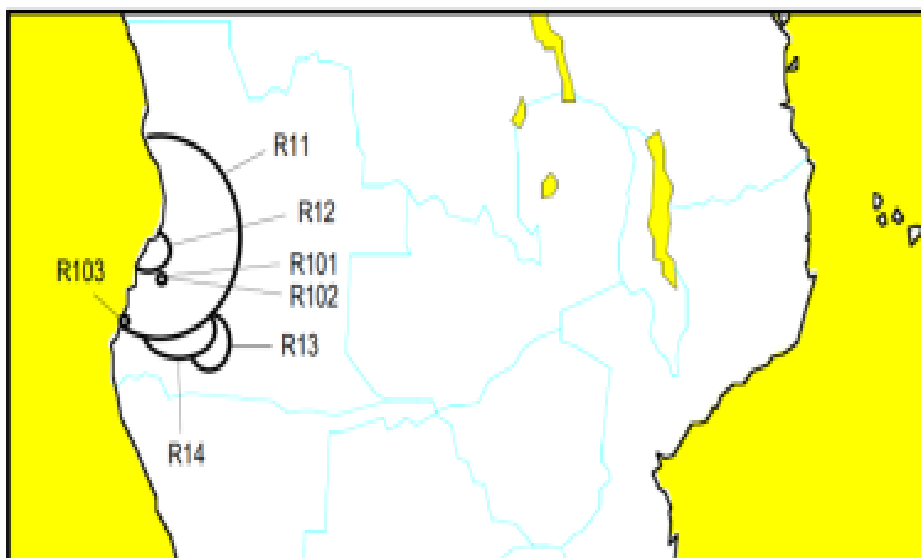
⁷ <http://coloquioslusofonia.blogspot.com/2012/01/angola-lei-sobre-estatuto-das-linguas.html>

⁸ <http://coloquioslusofonia.blogspot.com/2012/01/angola-lei-sobre-estatuto-das-linguas.html>

⁹ Guthrie’s classification of the Bantu languages is divided into 15 main zones according to their geographical location and language, categorized from Zone A to Zone S (Holden, 2002, p. 793), one of which is zone R corresponding to the wider Umbundu linguistic group with its subgroups. This classification was elaborated in his 1948’s *The classification of the Bantu languages*, published in London by the International African Institute. This classification has been used as reference for many *a posteriori* researchers who have updated it.

p. 14) and it has been one of the seven elected languages to be fully studied, experimented and, eventually, implemented and taught at public schools in Zones of its influence (Ponso, 2011, p. 9).

Figure 2: Umbundu group (R10)



Note. Map of the linguistic group as classified by the linguist Guthrie as in Maho (2009, p. 86)

Figure 2 shows the Umbundu Group (R10) as updated in Maho (2009, p. 86)¹⁰. It can be seen that the Umbundu Group is made of 7 ethnolinguistic subgroups, namely R101 (Kuvale), R102 (Kwisi, purportedly dead language), R103 (Mbali also Olumbali or Kimbari), R11 (Umbundu also, South Mbundu or Nano), R12 (Ndombe) R13 (Nyaneka) and R14 (Khumbi). The degree of distance between them also dictates the level of mutual intelligibility that they share. Umbundu (R11) is clearly the widest language of the larger Umbundu Group.

In Angola, like in all other African countries, Umbundu has to survive among other Bantu languages, many of which are very distant in terms of intelligibility. This is a corollary of what was briefly discussed in 1.2. above. In fact, the borders of nations (a brief discussion of the concept of nation will be included below) were violated so brutally that complete and well structured “nations” disappeared to the Portuguese overriding control. With this control came the cultural subjugation with its linguistic banning component (see 1.3.) reinforced by the famous Decree 77 of 9th December, 1921 of Norton de Matos (see 1.3.1.). Then, Umbundu had to be used in a territory made of different “nations” forced to be civilized by Portugal and whose language of instruction was

¹⁰ available at <http://goto.glocalnet.net/mahopapers/nuglonline.pdf>

peremptorily the Portuguese Language. This situation made Umbundu coexist with the other undervalued languages in a multilingual territory – Angola.

Figure 3: Umbundu in its multilingual context of Angola



Note. In the map above, the Umbundu language is presented as related with the other local languages spoken in Angola (Nurse & Tucker, 2001)

Figure 3¹¹ shows the Umbundu in the multilingual territory of Angola. The pink lines are the Angolan borders with other countries and the grey ones inside the country form the boundaries of the main language groups and subgroups. As it can be seen, in Angola, the Umbundu language community (in Group R, in lime green colour) coexists with the language communities in the Northern region of Angola (Group H, in beige colour) and the language communities in the Eastern part of the country (Group K, in mint colour).

As illustrated in the map, all these languages are reciprocally intelligible. Umbundu will most probably share more direct characteristics of intelligibility with Ndombe, Nkhumbi, Nyaneke, Zemba and Kwanyama within the Angolan territory, seen that they belong to the same group in Angola (R10). Nevertheless, the degree of intelligibility will be much reduced between Umbundu and the

¹¹https://www.sil.org/system/files/ reapdata/82/45/16/8245165576250264113481067886010116584/silesr2002_016.pdf.

languages of Group K, such as Nyemba, Mbwela, Nkangala, Luvala and Chokwe, on the one hand, and between Umbundu and the languages of Group H, on the other hand. Such languages are Nsongo, Sama, Bolo, Loanda Mbundu, and Mbangala as seen in the northern zone above Umbundu in the map of figure 2.

This section described Umbundu as a language existing within a multilingual country. This language has primarily been surrounded by other languages of the same group and those of both Group H in the North and Group K in the East. This contingency as a language in constant contact with other Angolan languages, and being the most spoken language of the country makes it necessary to be well studied before it is included in the Angolan education system.

One of the most difficult tasks that Angola has been facing since its independence is the lack of firm language policy and planning which can, without any hesitation, prevent languages from endangerment and/or decay. The following section, however, is an attempt to understand the nature of the Umbundu language maintenance during the time of the colonization.

1.6. The study of Umbundu in the colonial period

Although the colonial period was marked by the ban of native Angolan languages from schools and churches (see 1.3.1. above), there has been a great deal of production on local languages in the missions by individual missionaries of different Christian denominations before and even after the Norton de Matos's decree (Neto, 2012).

For the case of Umbundu, attempts to maintain it gives merit to Protestant missionaries, in particular. Their contributions have been directly linked to the study of the Umbundu language and culture, mainly in the second half of 1800. The earliest ever known research was offered by Rev. Sigismund Wilhelm Koelle (1823-1902), a German missionary, who published in the *Polyglotta Africana in 1854*, important information about the language of *Pangela* (how he called Benguela, an Umbundu-speaking province of Angola under the colonial period) (Fodor, 1977, p. 63; Koelle, 1854, pp. 84 –120). Another contributor to the study of Umbundu at earlier stages is from László Magyar (1818 – 1864), a Hungarian sailor, who presented, in 1859, a study of the ethnography and geography of the Umbundu people, a study which included important data about their language (Fodor, 1977, p. 13). Ten years after László Magyar's publication, in 1869, another researcher, Wilhelm Heinrich Immanuel Bleek (1827-1875) published the *Comparative Grammar of the South-African Languages* in London. This grammar, could very well explain the grammatical aspects of *O-*

tyi-hereró language which is reciprocally very close to Umbundu (Bleek, 1869, p. 104)¹². In fact, they belong to the same group R. As such they are intelligible and Umbundu's researchers have long taken advantage of works in this language to preserve theirs.

In the late 1800s an important research on Umbundu was published in Lisbon by Capello and Ivens in 1881. It was *De Benguella as terras de lacca*, in two volumes. Their work presents a vast number of used vocabularies of the Umbundu language, the language of the lacca. In addition to the aforementioned contributions, Reverends Sanders and Fay wrote *Vocabulary of the Umbundu Language, Comprising Umbundu-English and English-English*, in 1885, and, in the same year Stover published his *Observations upon the Grammatical Structure and Use of the Umbundu, or the Language of the Inhabitants of Bailundu and Bihe*. In 1887, Stover published an important book entitled *Otuikanda Tuokufetika Lokutanga Umbundu*, which was specifically designed to be a coursebook for the Umbundu language. Then, followed the work of J. Torrent, published in 1891, entitled *A Comparative Grammar of the South-African Bantu Languages* wherein he conglomerates languages from Zanzibar, Mozambique, The Zambezi, Kafirland, Benguela, Angola, The Kongo, The Ogowe, The Cameroon, and the like. José Pereira do Nascimento was the first native Portuguese who launched hands onto the Umbundu language with his *Grammatica do Umbundu ou Língua de Benguella*, in 1894. There may be many more contributions in the 1800, but these are the most influential works that could be considered.

The 1900s also produced significant contributions and publications on the Umbundu language. Héli Chatelain, a Swiss Protestant missionary trained in the United States of America, wrote “[A] few articles and papers on folklore and contributions to one book of hymns and parables in Umbundu, published posthumously in 1911” (Moser, 1983, p. 520). In 1934, an anthropologist and ethnologist called Wilfrid D. Hambly working for the Field Museum of Natural History of Chicago released a valuable work about the Umbundu language by analyzing the riddles, vocabulary, folklore, grammar aspects and so forth. The book is called *The Ovimbundu of Angola* (Hambly, 1934). In his turn, Albino Alves published in Lisbon, an Umbundu – Portuguese dictionary entitled *Dicionário Etimológico Bundo-Português* in 1951. José Francisco Valente, a member of the Catholic clergy produced, in 1964, a well-known grammar of the Umbundu language named *Gramática Umbundu: A Língua do Centro de Angola*. It was a valuable contribution for the study of the Umbundu language as it marks the first known public version of the Catholic intellectuals showing the way Umbundu

¹² This is a language of the same group as Umbundu (Group R) and very similar to it. For example, “o-n-dyuo-yatungwa” (and many other comparable grammatical constructions) in O-tyi-hereró language is found very similar to its counterpart in Umbundu Language which is “ondjo yatungwa”, meaning in both languages “the house is built”.

should be orthographically represented. As a matter of fact, there were some minor differences in the way Valente understood the writing of the language from the perspective of a person whose language of colonization was Portuguese.

Equally important is the work carried out by ethnographers which, although not being directly related to the development of the Umbundu Language *per se*, registers important linguistic information about this language and may be considered as part of the literature in Umbundu. To start with, a reference can be made to important contributions made by Marbel Stokey who wrote “*Vyovusenge*” (Things of the Bush) in 1916 (Soares e Agostinho, 2016, p. 503), “*Alivulu Akuala Olondaka Viwa kuenda Ovilinga Viovapostolo* (The Four Gospels and Acts of the Apostles)” a Portuguese-Umbundu edition, published in 1923 by the British and Foreign Bible Society of London (Soares and Agostinho, 2016, p. 503). In 1957, Westermann and Ward published a *Practical Phonetics for Students of African Languages* in London by the Oxford University Press and the International African Institute. Then, Lord Merlin W. Ennis wrote two important contributions that merit credit. The first is *Umbundu: Folk Tales from Angola*, published by Beacon Press in 1962 and, the second is *Embimbiliya Li Kola (The Bible, in Umbundu)*, published in Luanda, in 1963¹³. Gregoire Le Guennec and José Francisco Valente offered the public a bilingual dictionary entitled *Dicionário Português-Umbundu* in 1972. These are just some few illustrative cases of what has been produced in Umbundu.

All in all, important bibliography made on behalf of the development of the Umbundu language as shown above indicates that there has been a constant effort to maintain it. It can be said that in the second half of the 1800s more serious research work related to the understanding and development of the Umbundu language was decisively launched and this has helped much to create a sustainable background for the challenges of today. In the 1900s, there was a very timid production, if the development of technology and scientific is taken into account as time went by. In this period, there was a clear combat against the development of native languages, mainly with the entrance of the Norton de Matos’ prohibitions by law in 1921. Moreover, the prohibition of the use of native languages in Angola has contributed to the degradation of necessary positive attitude towards Umbundu and other languages.

The following chapter has been designed to revise some important concepts related to language planning and policy by differentiating the two terms. The types of language planning, approaches to language planning, language maintenance, language shift and language revival, the

¹³ More detailed data can be found at <https://www.jw.org/en/publications/bible/study-bible/appendix-c/divine-name-new-testament-2/>.

ecology of language and language and education, the insertion of national languages, and the like will be discussed in it. A difference will be established of bilingual education and literacy, multilingual education and literacy. Then, the types of orthographic systems and the choice of writing systems will be discussed as well.

2. CORPUS PLANNING WITHIN LANGUAGE POLICY AND PLANNING

The previous chapter provided a background of how, among other countries, Angola has emerged out of the decisions established by the Berlin Conference (1884 – 1885). It also looked at the way the resulting new owners of the southern region of Africa (namely, England, France, Germany and Portugal) used their national languages as real symbols of hegemony and exclusion of African at all levels possible in contexts where different language communities experienced the ban of their languages, or, at the least, the devaluing of their languages. As consequence, this policy has prevailed in many countries, including Angola. Unlike the preceding, this chapter has been conceived to address language corpus planning in a way that it can serve as a referential background to the understanding of this research project, namely with regard to the orthographic harmonization of languages towards their insertion into education systems. It is here important to consider that corpus planning is a co-part of a wider set of language reference and planning. As such, a differentiation between language policies and language planning will be briefly revisited in the beginning sections of this chapter followed by a discussion of corpus planning as a dimension of language planning.

2.1. Language policy and planning: An essential distinction

A significant number of well-established academics in the field of Sociolinguistics, Applied Linguistics and Educational Linguistics, have devoted their time to understand the process of language “engineering” (for example, Holmes, 2008; Hornberger, 2006; Madiba, 1999; Cluver, 1992; Haugen, 1987, 1972,; van den Bergh, 1968; Haugen, 1966a; just to mention some). Thus, two crucial processes have to be discussed and clarified here. The first is related to language policy and the second has to do with language planning. Although they are two related concepts – one cannot go without the other (Jones, 2015, p. xiii; Davies & Ziegler, 2015, p. 2; Darquennes, 2013, p. 106; Hornberger, 2006, p. 25 and) – it is important to clarify that the two terms can be used complementarily. This way, it is believed, it will be easier to understand the reason why they cannot go apart. The following section will devote space to discuss language policy and language planning.

2.1.1. Language policy

Language policy (also called language management) refers to the process which happens at the political level involving budgeting and managing in the parliamentary level, and in other institutions

under government control as well as non-governmental agencies (see Hartshorne, 1987; Heine, 1992) which use processes that establish merit and guiding principles allowing, at the same time, for ongoing needs for punctual changes according to varied contexts of language use. Some contextual domains of language use, under the umbrella of language policy, comprehend the defense area, education, teaching additional languages, religion, sign language and society at large, including the family (Spolsky, 2012, pp. 9 – 13). Depending on political context, at this stage of language management, patriotism and the willing to commune with others, under a common sense paradigm, may, *de jure*, flourish to supplant the spirit of separation, segregation or exclusion when deciding on language(s) which people of a unitary territory, living in different geographical points, have to acquire, learn and use formally or informally (Siiner, Hult, and Kupisch, 2018, p. 1). The patriotic sentiment does not necessarily need to be the first thing to occur (Darquennes, 2013, p. 106). Again, on behalf of the feelings mentioned here, a class of people – habitually, people governing the territory – may decide to impose the direction that language policies shall take according to an agenda which may aim at segregating people for better control and domain (Ngcobo, 2007, p. 7; Cobbarrubies and Fishman, 1983, p. 13; Lanham, 1978, p. 21). In South Africa, for example, “rule and divide” systems had in mind a language policy which led to intense conflicts and “caused resentment from the indigenous people and resulted in the Soweto youth demonstrations against Afrikaans in 1976” (Ngcobo, 2007, p. 7).

The choice of languages for a country, mainly in the post-independence era may lead decision-makers to elect among endogenous languages, the language of the colonizer, (e.g. the in Angola and Mozambique) or a mixture of both endogenous and exogenous languages to play roles as official languages as in South Africa (see Ngcobo, 2007, p. 7; Potgieter & Anthonissen, 2017, p. 131). If an autochthonous language is chosen, the process will be called endoglossic language policy. On the other hand, if the language of the colonizer is considered to be the official language, it will be called exoglossic language policy. The following section sheds some light on these two crucial processes.

An important aspect to be taken into account when analyzing the process of language policy is that it can be categorized into two principal types, mainly with respect to the African context (Heine, 1992, p. 23). One is the endoglossic kind of language policy which occurs when political decisions related to the choice of languages promote languages of African origin to be the principal agencies of communication and instruction in a country, in all vital levels of life. These countries are called endoglossic nations. For example, Tanzania uses Swahili (Kiswahili), Somalia has elected Somali, Sudan communicates through Arabic, Ethiopia utilises Amharic and Guinea uses eight languages,

namely “Fula, Manding, Susu, kisi, Kpelle, Loma, Basari and Koniagi” (Heine, 1992, p. 24). In endoglossic nations, the kind of language policy that pursues the total emancipation from the Western languages is said to be *active*. In its turn, a *passive* endoglossic language policy happens in countries, such as Botswana, Burundi, Lesotho, Malawi, where Tswana (or Setswana), the Rundi or Kirundi, the Sotho or Sesotho and Chichewa and Nyanja, respectively, are promoted but they are still critically dependent on exogenous colonial languages for the most important issues of their everyday life (in education, administration, government, just to give a few examples).

Unlike endoglossic language policies, there are those ones which have been thought to endorse the colonial language officialization policies. This kind of language policy is called exoglossic language policy. In fact, most of the countries in Africa, including Angola, Zimbabwe, Niger, Benin (just to mention a few) have followed this approach after they proclaimed independence (Heine, 1992, p. 25). For these countries, the use of African languages has not guaranteed enough legal weight, and most often, if not always, the languages are not “national” as such and are not used as means of instruction. Moreover, exoglossic countries such as Kenya, Mali, Senegal and Uganda fall into the category of language policy where the legislators acknowledge the existence of *de facto* “national” language implemented as such in the country (Erastus and Erastus, 2013). In fact, countries such as Nigeria, Togo and Niger fall into the category of those ones where the regional languages are even taught at schools and used for official affairs to visible extents (See Ihejirika, 2017; Okonkwo, 2016; Danladi, 2013; Heine, 1992).

Language policy is part of a complex process of societal language treatment. In embarking into such process, important factors should be taken into account, i.e. the level of education and scientific awareness of the people (agency) who are to decide over language choice for educational functions should be taken very seriously (McEntee-Atalianis, 2016; Ricento, 2006; Baldauf, 2004; Ricento, 2000; Ricento & Hornberger, 1996), as language policy, when enshrined in the law, constitutes a mandatory provision. In fact, McEntee-Atalianis (2016, p. 5) has categorized three such groups influencing language policy processes: (1) Experts – those who are socially well-positioned and respected, such as language researchers and/or those in experts who have been intensely interested in language policy, (2) Prestigious citizens such as respected writers, eloquent entrepreneurs; church leaders, doctors of law negotiators and famous people, and (3) politicians, - “including language planning officials – ”. As such, the multiplicity of agencies may lead to sociolinguistic conflicts of unexpected consequences because in general, those who are specialists may not be in the political arena to make decisions with the potential to become laws based on

research. On the other hand, people with prestige in a society may influence law making and enforcement even though they do not have scientific information about the language issues going on in their society. Those who constitute the lawmakers working at the political level, maybe the most obstinate to accept view points of specialists for they generally decide about issues without the specialized knowledge and skill to understand coherent. As it will be seen below, the Umbundu orthography was developed by missionaries, most of which did not have expert preparation to deal with the orthographies of languages that they did not know very well. The multiplicity of agencies may make the decisions about languages status and their use into the education systems somewhat difficult.

2.1.2. Language planning

Language planning is often a subsequent stage of language policy, and many researchers have dedicated significant time to it, especially with the focus on inclusive education (Kamwangamalu, 2011; Kaplan, 2011; Kheng, and Baldauf Jr, 2011a; Kheng and Baldauf Jr, 2011b Zhao, 2011 Ager, 2005; Liddicoat, 2005; May, 2005; van Els, 2005). According to Reagan (2010, p. 31), “Language planning is both an academic discipline and a practical activity resulting in the development and implementation of a specific language policy”. It can be understood as a collective supply within a social planning structure, and a “decision-making” operation aimed to work out problems of communication [through both written and oral modalities, *my emphasis*] (Daoust, 1997, p. 439). In fact, it can be seen that language planning corresponds to an endorsement that leads a government or any person or persons to decide on the allowance to influence future change of a language in and for use, both functionally and structurally (cf. Reagan, 2010, p. 33; Kaplan and Baldauf, 1997, p. 3; Cooper, 1989, p. 45,) to solve societal conflicts or ameliorate educational programs and curricula (See UNISA, 2011, p. 14). Conflicts resulting from questioning how a word is correctly spelt and pronounced, what its etymological meaning should be, or what orthographical system is at stake, will need language planners to explain and sort them out (Reagan, 2010, p. 34). In that context, a construct of language planning should explain the necessary political merits encompassing inclusive initiatives so as to give space for all languages to be protected, cultivated and adjusted on behalf of the peoples who use them, i.e. the process should be democratic (Phillipson, 1992, p. 86). For Language planning to be accomplished there are important dimensions that should be taken into account (Kloss, 1969). The section that follows will be devoted to describing them.

2.1.2.1. The four dimensions of language planning

At the beginning of the studies on language planning there used to be two important dimensions of Language Planning: *corpus* planning and *status* planning (Kloss, 1969, p. 81). Then, language planning included one more dimension – *acquisition planning* to make three fundamental components (Wright, 2012, p. 64; Coulmas, 2005, p. 186; Crystal, 2003, p. 366; Daoust, 1997, p. 449; Wiley, 1996, pp. 107-16; Cooper, 1989, p. 33;). According to Darquennes (2013, p. 106) and Baldauf (2004, p. 3), four such dimensions can be mentioned. To start with, they consider *corpus planning*, which comprehends the *codification* of the norm, standardization procedures (i.e. graphization, grammar and lexis), as well as the *elaboration* at the functional level (i.e. terminological modernization) and stylistic development. The other dimension is *status planning*, which encompasses the *selection* of the norm, decision procedures, identification of a problem and allocation of norms, and it encompasses, in its turn, the *implementation* in education, the adjustment procedures and the stage of evaluation. There also is the *acquisition planning* which is related to the “decisions concerning the teaching and use of language, and their careful formulation by those empowered to do so, for guidance of others” (Wiley, 1996, p. 109). The fourth dimension to consider here is that of *image or prestige* - “actions that aim at promoting the acquisition of a language (variety) and, by doing so, increasing its number of users” (Darquennes (2013, p. 106). *Image (prestige)* may be linked to *ethnic* or *civic identity* “as in the case of Québec” where *image* serves as a *method* of putting language into practice and influencing its direction (Baldauf, 2004, p. 4). *Image* also is linked to the *raison d’être* of planners and what they themselves do taking into account the point of view of the community, “as in Malaysia” (Baldauf, 2004, p. 5). As it can be perceived these components or dimensions are closely related to each other within the circuit of language planning. Table 1 below represents a synopsis of the dimensions of Language Planning as in Baldauf (2012: 248).

Table 1: A developmental model for language planning goals

Approaches Types (overt – covert)	1. Policy Planning (on form) Goals	2. Cultivation Planning (on function) Goals
1. Status Planning (about society)	<i>Status Standardization</i> <ul style="list-style-type: none"> ● Officialization ● Nationalization ● Proscription¹⁴ 	<i>Status Planning</i> <ul style="list-style-type: none"> ● Restoration ● Revitalization ● Reversal Maintenance Interlingual Communication <ul style="list-style-type: none"> ● International ● Intra-national Spread
2. Corpus Planning (about language)	<i>Standardization</i> Corpus <ul style="list-style-type: none"> ● Graphization ● Grammatication ● Lexication Auxiliary Code <ul style="list-style-type: none"> ● Graphization ● Grammatication ● Lexication 	<i>Corpus Elaboration</i> Lexical Modernization Stylistic Modernization Renovation <ul style="list-style-type: none"> ● Purification ● Reform ● Stylistic simplification ● Terminological unification Internationalization
3. Language-in- Education Planning (about learning)	<i>Policy Development</i> <ul style="list-style-type: none"> ● Access Policy ● Personnel Policy ● Curriculum Policy ● Methods & Materials Policy ● Resourcing Policy ● Community Policy ● Evaluation Policy 	<i>Acquisition Planning</i> Reacquisition Maintenance Foreign/Second Language Shift
4. Prestige Planning (about image)	<i>Language Promotion</i> <ul style="list-style-type: none"> ● Official/Government ● Institutional ● Pressure group ● Individual 	<i>Intellectualization</i> Language of Science Language of Professions Language of High Culture

¹⁴ Proscription occurs when there is a firm decision, generally, official one, for banning a language. An example of proscription is the banning of autochthonous languages of Angola from schools and missions by the High Commissioner of Angola Norton de Matos as discussed in the previous chapter.

Table 1 shows a developmental model for language planning goals. The understanding of these dimensions is of paramount importance to this study in that they inform about what corpus planning encompasses and how important it is to carry out a corpus planning for the harmonization of orthographies of the Umbundu language for educational purposes in the Angolan education system. One way of preserving the cultural history of a given community is also by understanding how the oral fonts are registered in a way that the future readers take the legacy. Corpus planning constitutes a more focused way and opportunity of making an endangered language such as Umbundu revive, as discussed in subsections 1.2 and 1.3 above regarding the famous Decree 77 of the Governor José Maria Mendes Ribeiro Norton de Matos who in 1921 signed for the banning of languages such as Umbundu, putting them in a dangerous position of disappearance. Graphization, grammatication and lexication of the Umbundu language constitute the basic work for a common writing system to exist. Lexical and stylistic modernization of this language promises to engenders renovation at the levels of Purification, reform, stylistic simplification, terminological unification and the necessary internationalization of the language.

In sum, language policy and language planning go hand in hand throughout this complex process of language management for achieving social justice (see Hornberger, 2006). As advised by Fettes (1997, p. 14), language planning and language policy should be connected to substantiate the decisive assessment of language policy with language planning playing the role of provision for the guarantee of serious undertaking, while language policy plays the role of “testing these ideas against actual practice in order to promote the development of [...] language planning models” at the governmental sphere. As a matter of support, Taylor-Leech (2016,) align with Nekvapil and Sherman (2015) in the sense that language planning and language policy are not separate entities, for they conglomerate a plethora of research branches which meet at the establishment of order and control of the existing language-related issues and their role in the communicative acts.

The preceding section has dealt with the developmental models of language planning and its components. Corpus planning is of crucial importance for this study given that it deals with graphization, grammatication and lexication. As it will be discussed in Chapter Six, the various problems caused by the development of the Bantu languages orthographical systems by non-native researchers has led to the current orthographical standards of the Umbundu language based on the European's language structures (Werner, 1919). So, this model considers the agents of language planning whose roles deserve attention in the following subsection.

2.1.2.2. Agents of language planning

Language planning is an endeavour carried out by clearly defined agents (Cooper, 1989 & Kaplan & Baldauf, 1997). For example, Serralvo (2013, p. 26 ff) has presented a number of agents classified according to the following perspectives:

a. From the top-down perspective, the most important agent is the government by means of its specific State departments with legal and executive power, such as the local administration, education sector and the State media.

b. From the bottom-up perspective, the agency is led by organizations of the civil society and/or NGOs, individual initiatives. As Cooper (1989, pp. 183 – 4) puts it, “writers, poets, linguists, language teachers, lexicographers, translators...missionaries” can be influent agents of language planning.

So, language planning actors may play roles from different segments of the society. Both the elites and the counterelites work towards the preservation of their groups’ interests, and the communication they engage in may help reach agreements that can benefit both sides.

2.1.2.3. Factors that determine language planning

There are important factors determining language planning. As discussed in Serralvo (2013), leaning based on the contributions given by Cooper (1989) those factors can be interpreted as follows:

a. Sociopolitical and economical circumstances (“situational factors”). These factors correspond to moments tendentiously unstable, including economical crises, natural catastrophes, technological innovations and the like that may influence decision making by those who hold power in order to appease the community insurgency to adhere to governmental policies. Cooper (1989, p. 93) illustrates this factor with the famous race riots that happened in Soweto, South Africa to protest and force the *Apartheid* regime to relinquish its intention to mandate Afrikaans as medium of instruction for African community’s schools.

b. Permanent States of affairs (“structural factors”). These factors comprehend relatively permanent “political, economical, social, demographic, and ecological” shapings of the society (Cooper, 1989, P. 93). The case of the governmental decisions that have to be made on the conciliation of two orthographical systems of Umbundu for the education system can serve as example of permanent States of affairs on language planning (Jimbi, 2018).

c. Societal habits (“cultural factors”). These factors are determined by the attitudes and ideals of the members of a given community or those of its subgroups. These include norms and principles to be observed in different circumstances, ideologies, roles of the members of society and the like. An illustration for these factors can be the different attitudes exhibited by Catholic members, on the one hand and Protestant members, on the other, towards the Umbundu orthography (Jimbi, 2018).

d. Exogenous conditioning (“Environment factors”) (Cooper, 1989, p. 93). These factors are the counterparts of the previously presented factors with the difference that they “influence decisions” from outside the government, such as the “international political environment...international agreements, obligations, and pressures (Cooper, 1989, p. 93). To illustrate exogenous conditioning as an important factor of language planning, the cooperation between the South African The Centre for Advanced studies of African Society (CASAS) and the Angolan *Instituto de Línguas Nacionais do Ministério da Cultura* for the harmonization of the orthography of the Angolan *Bantu* languages (including Umbundu) can be taken as example (Pedro *et al.* 2013).

e. Bibliography (Information). This factor has to do with the availability and gathering of necessary information to sustain decent decision-making. The process of data gathering demands capacity to choose relevant data among the millions available for processing, and the awareness of the consequences the data may bring to the organizations. In the case of Umbundu, there has still been very little contribution both from the State and the individual agencies and the few existing findings, so far, “are largely unfounded and even self-contradictory” (Schadeberg, 1982; Pedro *et al.*, 2013).

The above factors are of relevance in language planning – mainly with regard to corpus planning at the level of orthography applicable to the study of the Umbundu language – given that they have influence in the process of language related decision-making by different kinds of agents, as language planning is a social endeavour which demands cautiousness on the part of the players regarding what interferences there are that need further studies and considerations for productive results. In that case, no single factor can be said to be more adequate than the other in the process of language planning. Each case is a case demanding analysis of a given factor as influencing the process of language planning.

2.1.2.4. Important principles to consider in language planning

In the process of language planning, there are principles that should be considered by the practitioners so as to guide themselves towards expected results. These principles have the potential

to help language planners to be aware of the hindrances to overcome and facilities that there are to take advantage of. Serralvo (2013, p. 32) has enlisted these principles as quoted below:

- The convergence principle: the impact of regulatory actions will be greater so far as factors can favour one same option. The diversity of proposals weakens the action.
- The dominance principle: the impact of the socially dominant actors will be greater to that of the non-dominant actors.
- The persistence or the social inertia principle: the social and linguistic behaviour has a tendency to reproduce itself; change is more difficult than innovation in contexts where scientific research tends to be taken for granted.
- The linguistic system's internal coherence principle: although there is intrinsic change in every language, as a system it has its own internal coherence which repels some options of change spontaneously.

The principles above are of relevance for this study given that decisions about the orthographical model to be used depend, to a greater extent, on the correlation of influences existing among the above principles. It takes time, support and commitment to reach consensus on which orthographical standard should be applicable, mainly when different traditions are put to test of credibility.

2.1.2.5. Corpus planning as a dimension of language planning

Corpus planning as a dimension of language planning deals, here, with language usage, given its role in the process of information production, information transfer, information consumption and information preservation through printed and reading materials in diverse institutional formats (Ndimande-Hlongwa, 2010, p. 208; Olohan, 2004, p. 1;). It is a “primarily linguistic” process (Goundar, 2017, p. 86; Hill, 2010, p. 46; Lo Bianco, 2010, p. 146; Ndimande-Hlongwa, 2010, p. 207), one which deals directly with the recommendation of an “ideal form” of writing and speaking to be followed by the members of the society, “while everything outside of the standard is often rejected and considered an ‘error’ or ‘dialect’” (Odendaal, 2013, p. 186; Mangoya, 2009, p. 26). In fact, from the policy planning standpoint, corpus planning focuses on the coding of a language, including the standardization of words and expressions, the proposal of auxiliary codes (such as sign language) and the definition of how utterances should be put into the written mode (Leimgruber, 2019, p. 80; Darquennes, 2013, p. 106; Baldauf, 2004, p. 8). From the functional point of view (i.e., cultivation), corpus planning deals with the demands of languages in a way that previous technological means could not do, at both lexical and stylistic levels (Baldauf, 2012, p. 248), renovate the language forms to make them more apt for the environment they are used, reformed and simple in style, unify the terminology and internationalize (Liddicoat, 2005, p. 993 ff).

The following subsections – starting with “Standardization of corpus” – will be devoted to the revision of previous studies on corpus planning, with necessary details of its influencing components (codification, graphization, Lexication and so forth), in a way that they can elucidate the understanding of the complexities of the processes involved in the standardization of a language, and then be used as a basis to comprehend the phenomena in the context of this study.

2.1.2.6. Tasks of codification: The standardization of *corpus*

Codification is a subcategory of corpus planning which deals with the technical treatment and validation of a selected variety of language to be standardized, and it constitutes the cornerstone for the effectiveness of the whole process of language planning (Vandenbussche, 2015, p. 6; Baldauf, 2012, p. 234). As such, it is a process which is heavily prone to sociopolitical and economical interference (Liddicoat, 2005, pp. 944 – 955), because decision makers (e.g. religious elite, political elite, academic elite established and conformed with the already existing norms) who control the mechanisms of sponsorship for such kinds of endeavour may either not have enough awareness of the importance of this kind of endeavour or, even worse, become the fiercest opponents of those who decide to engage in it (see Garvin¹⁵, 1993, pp. 45-48). Standardization of *corpus* is an important task in the codification process. And, their main components - graphization, grammatication and lexication – will be discussed in details in 2.1.2.7. and 2.1.2.8., below.

Basically – as Stated above - standardization refers to the processes of normalization (Cf. Linn, Sandden & Piekkari, 2018, p. 22) of linguistic items of a selected language or its variety, frequently under a “language’s common core” (c.f. Elkarte, 2010, p. 17). So, a discussion about “standardization of corpus” in this study will entail an understanding of the evolution of the concept [of standardization] as applied to language planning. *Corpus*, here, must be understood as any utterable linguistic item susceptible of being represented in the written form in the way McEnery and Wilson, (2001, p. 2) and McEnery and Hardie (2011, p. 3) take it¹⁶, including “Sign languages”, which are frequently disregarded in important scientific research related to language standardization (Ghyselen, Delarue & Lybaert, 2016, p. 86).

¹⁵ Garvin (1993) discusses the functions of standard language and the resulting attitudes of the user. Garvin reports about five functions of standard language, namely unifying, separatist, prestige, participatory and objective frame-of reference. As to the attitude these functions are linked to the first two functions have relation to loyalty, the third is linked to community pride, the fourth is linked to the willing to modernize the language and the last one is relates to the consciousness of the adopted norm.

¹⁶ These authors consider that the notion of corpus encompasses systematically arranged formal linguistic components, such as morphemes, words, syntactic patterns which mirror various functions– be they semantic or pragmatic. Gries (2009, p. 8) believes that “The texts that make up the corpus must have been produced in a natural communicative setting. That means that the texts were spoken or written for some authentic communicative purpose”

Again, the concept of standardization of corpus and the way it has been interpreted by specialists has undergone important updates, thus far, as more research on it started to emerge, with time, to capture more attention and interest, mainly on the part of sociolinguists (Darquennes, 2013, p. 106; Liddicoat, 2005). To start with, Haugen's four-staged model of corpus standardization (sequentially organized as *selection*, *codification*, *implementation* and *elaboration*) (Deumert & Vandebussche, 2003b, pp. 4 - 9) has been exceptionally influential among sociolinguists, mainly for its recognized possibility to be applied in comparative terms with other models of standardization (Coupland & Kristiansen, 2011, p. 20, for more detailed information). However, this model has proven to be inadequate for generalizations today. In their turn, Milroy and Milroy (1985a, 1985b, 2012; Pillière & Lewis, 2018) have come up with a proposal of 7 stages, namely selection, acceptance, diffusion, maintenance, elaboration of function, codification and prescription. A crucial point they make of the difference from the other model that the stages of language standardization do not need to be sequential, and written and spoken texts need to undergo similar stages simultaneously. Also, it has been seen that Haugen's model cannot respond, for example, to the fact that there have been, in the planet, many crossborder languages which naturally face the challenge of being standardized in differing ways (for the imperative of existing in and being influenced by two or more different sociopolitical, economical, cultural and/or religious contexts), despite their belonging to the same "diasystem" (Darquennes & Vandebussche, 2015, p. 7). In that case, a language or a language variety may take either a "*monocentric*" standardization, i.e. when a language or a language variety consists of a unique linguistic norm to be observed by everyone in a certain space and time, or "*polycentric*", that is, when a language or a language variety comes to be standardized in multiple ways in a certain time and place (Stewart, 1968, p. 534). This is a demonstration of how difficultly Haugen's top-down model would explain the current differing motivational factors driving the agents into the standardization process, as well as the description of possible hidden goals during the process. As if not enough, nowadays many questionings of authority come from different autonomous agents, including linguistic rights activists, the academia, and the members of communities at large (Coupland & Kristiansen 2011).

Many researchers have considered the limitations of Haugen's model (Haugen 1987b, p. 63) and have offered alternative explanation of corpus standardization process. For example, Neustupný believes that 'any act of language planning should start with the consideration of language problems as they appear in discourse, and the planning should not be considered complete until the removal of the problems is implemented in discourse' (Neustupný, 1994, p. 50). In the mind of some

researchers corpus planning corresponds to aspects which are essentially linguistic, i.e., aspects appertaining to language, namely innovations in orthography (Kaplan & Baldauf 1997, p. 38). For Cooper (1989, p. 125 ff), corpus planning is a process involving graphization, standardization, modernization and renovation.

With the ideas above in mind, some minority languages in Europe, including Cornish, Corsican and Galician, were studied with researchers being provided with freedom to choose how they would go about it individually (Darquennes & Vandebussche, 2015, p. 8) Also, the appearance of regionalist activism in the 70s was one of the factors attracting political attention towards language issues. Another important factor was the direct intervention of the European Union through the European Charter for Regional or Minority Languages in issues related to the promotion of minority languages in the 80s. In this regard, corpus planning was specifically addressed as Stated in Darquennes and Vandebussche (2015, p. 5) that

When it comes to corpus planning measures for minority languages, the Charter briefly mentions activities that (i) support the role of the language in the media and the courts and (ii) aim at the maintenance and development of administrative, commercial, economical, social, technical or legal terminology (Art. 9 and 12). Those measures that are directed at the training of minority language teachers and the availability of minority language education (i.e. acquisition planning) also imply (a concern for) actions that aim at modifying the corpus of the minority language.

This sparked the appearance of many more language rights-related groups, academies, institutes and lobby groups to introduce linguistic activism with a focus on transversal study, revival and maintenance of minority languages, independent of the size of the linguistic community at stake.

The aforementioned study has also demonstrated that the process of standardization can be directly affected by the political, social, economical and cultural factors in the context they take place. For example, French is a majority language in the entire France, save for its being a minority one in the Val d'Aoste province in Italy. The same is true of Dutch which is a majority language in The Netherlands, although it is a minority one in Frans-Vlaanderen region in France. This aspect explains the need to consider both top-down and bottom-up approaches to language standardization as diverse agents give their contribution to the process itself independent of the official authorization for such enterprise. As Darquennes and Vandebussche (2015, p. 9) put it, the process of standardization may proceed in a way that

Codification precedes selection or elaboration precedes implementation.[...] It concerns a jumble of competing top-down as well as bottom-up standardization processes, instead. That different sorts of highly or less visible and influential language planning actors [...] instigate these processes, adds to the complexity of the case.

The above discussion has been demonstrated, for example, in the research work carried out by Vosters and Villa (2015) who also recognises that overcoming the impact caused by fully-supported language standardizers from the top social positions is not an easy task. However, the debate about the direction that language standardization should take today is so democratic that questioning the authority of language planners, even those from the elite, is so frequent that it helps impeding the imposition of non-agreed standards (Coupland and Kristiansen 2011, p. 27). Even if politically imposed, those who do not agree may feel themselves as the defeated part of the process of standardization, and influence the members of the community who master the language and believe in them not to adhere the implementation of the standard. This makes it necessary for considering Garvin's model of language planning which moves under three important questions: "what is a standard language? How does a standard language serve its users?", and also "what are the conditions required for the development of a standard language?" (Ndimande-Hlongwa, 2010, p. 211). Garvin's model is functional for the African multilingual context in that it allows for flexibility in codification and "intellectualisation" (*ibid.*).

All things considered, an important lesson can be learnt that although Haugen's four-step model of language standardization remains crucial as a comparative and resourceful instrument for language planners, it is no longer capable of explaining all the interdependent mechanisms involved in language standardization processes that are on demand today. The standardization of minority languages everywhere demands more than just prescribing a model projected in a top-down manner. It demands, it can be reclaimed, the understanding of all the factors (political, economical, cultural, individual, etc.) related to the concrete environment of a studied language or language variety, so as to approach the whole process of each minority language standardization as a unique one, only approachable on the basis of the various possibilities conducive to a thoroughly pondered and agreed decision prior to its codification as a finished product.

As aforementioned, for a process of standardization of corpus to take course, it has to respond to specific conditions, including that it has to undergo the processes of graphization, grammatication and lexication. The following subsections will be reserved for detailed accounts of each of these requisites starting with graphization.

2.1.2.7. Graphization, grammatication and lexication

One important element of corpus standardization is the process of graphization. Many researchers have defined graphization as the task of deciding which writing system shall be adopted as well as

how the modification of orthography shall be approached (Liddicoat, 2005; Ndimande-Hlongwa, 2010). As to the writing systems to be selected, the language standardizers tend to choose from three most prevailing arrangements – the alphabetical (e.g. Bambara language in West Africa) the syllabary (e.g. Vai and Somali in West and East Africa, respectively) and the logographical (e.g. the Chinese writing system) - or a combination of them according to the perceived needs in the context of the specific language planning (Liddicoat, 2005, p. 995). The first arrangement is related to the long-desired sound-letter correspondence while the second one is related to the representation of syllables by means of symbols. The third arrangement entails the representation of a unit of meaning by one symbol. In Angola, the writing system of regional languages, including Umbundu (Pedro *et al.*, 2013) is alphabetical brought by the colonial potency – Portugal.

The use of an existing system (for example, an orthographical one) will be facilitated by the support of a vast range of resources available. The main concern of this choice is that there always is lack of effective sound-to-letter representativeness, which often causes forcing the language to be unnecessarily corrupted insofar as the representation of natively produced phonemes is concerned. For example, Jimbi and Sicala (2020: 128) have traced some of the consequences that the lack of sufficient alphabetical letters to respond to the demands of the sound system of the Umbundu language could cause as follows:

- (i) The confusion in pronouncing the letter <c> when followed by the vowels a, o and u. e.g., the proper noun *Cokwe* vs *Čokwe*¹⁷ vs *Chokwe* vs *Tchokwe* ['kɔkwɛ] [tʃɔkwɛ] [ʃɔkwɛ] [tʃɔkwɛ]
- (ii) The confusion with the pronunciation of the letter < s > between vowels . For example the proper noun “*Sasoma*” may be pronounced as [sa 'zɔma] for a Portuguese literate instead of Saçoma¹⁸[sa 'sɔma]
- (iii) Double representation of the same sound. For example, Kuando Kubangu (Kwando Kuvangu) vs Cuando Cubango; Kwanza-Sul vs Cuanza-Sul

¹⁷ The <č> exists in the Czech writing system with the sound value of [tʃ] which the Portuguese alphabet does not offer, having given to the forcing of <c> of Portuguese to represent the sound [tʃ] in the first word – Cokwe – above for it is a local language word. Jimbi and Sicala (2020) have found <č> innovative and functional – under the principles of consistence, simplicity and economy – for the Umbundu orthography, and so capture the sound [tʃ]. For the learner and the common user, this may cause serious problems. And, as time passes the correct way of pronouncing the native words undergoes progressive corruption from generation to generation.

¹⁸ Jimbi & Sicala (2020, *loc. cit.*) propose the use of <ç> as a letter whose sound representation cannot change for the position it takes in a word, giving its consistence for the reader.

- (iv) The unnecessary use of trigraphs, digraphs and inappropriate diacritics. For example, ng' (PEDRO et al., 2013), ñg (CHIMBINDA, 2015) and ñ (DANIEL, 2002) for the phoneme [ŋ], when Valente (1964) suggests a more economical use of <g̃> as alternative.
- (v) The imposition of the sound [r] in the Umbundu phonology and phonetics. For example, *Chongoroi* instead of **Tchongoloi** [Čongoloi, as proposed by the authors] *Onjiri* (sable) for *Ondjili/Onjili* [onj̃ili, as proposed by Jimbi and Sicala (2020)].

Another aspect of standardization is called grammatication. It has to do with the characterization of the rules governing the parts of words and their combinations in texts. When standardizers decide on grammatication, they mean that a focus will be put on the structural aspects of words and expressions constructed out of the combination of characters that have been approved as information highlighters – “*Questions, Instructions, Advices, Request, Information, Warning and Intention*” together with the necessary punctuation marks - to simplify and support the understanding and learning of the standard language (Kaplan & Baldauf, Jr., 2003, p. 212). Agreeing with the postulate above is Liddicoat (2005, p. 998), who takes grammatication as the “reduction of the social and/or regional variation found in the existing spoken language ecology and formulating a set of [...] rules” to be used as prototype for writing various needed formal documents, including school manuals. The grammatical model of the standard language is often proposed or “imposed” by a well-established social, economical, political or academic group, or it is one that has established itself historically.

For the standard grammar to be a prescribed reference for the society, it has to undergo a process of description by specialists (cf. Liddicoat, 2005, *loc. cit.*). When it has been fully codified, the grammar will be established as an official reference which marks the difference from other concurrent standards of the same language. These other standards may be declared as “error” versions of the language, or simply labelled as dialect. In Angola, for example, there are not officially established grammars or dictionaries of the Umbundu language. Instead, individual initiatives, supported by groups, have been proposed (Do Nascimento, 1894; Valente, 1964; Daniel, 2002; Malumbu, 2007, Le Guennec & Valente, 2010; Daniel, 2010) and users, indiscriminately, chose one option (generally the one used in the Bible of their church) and proceed accordingly.

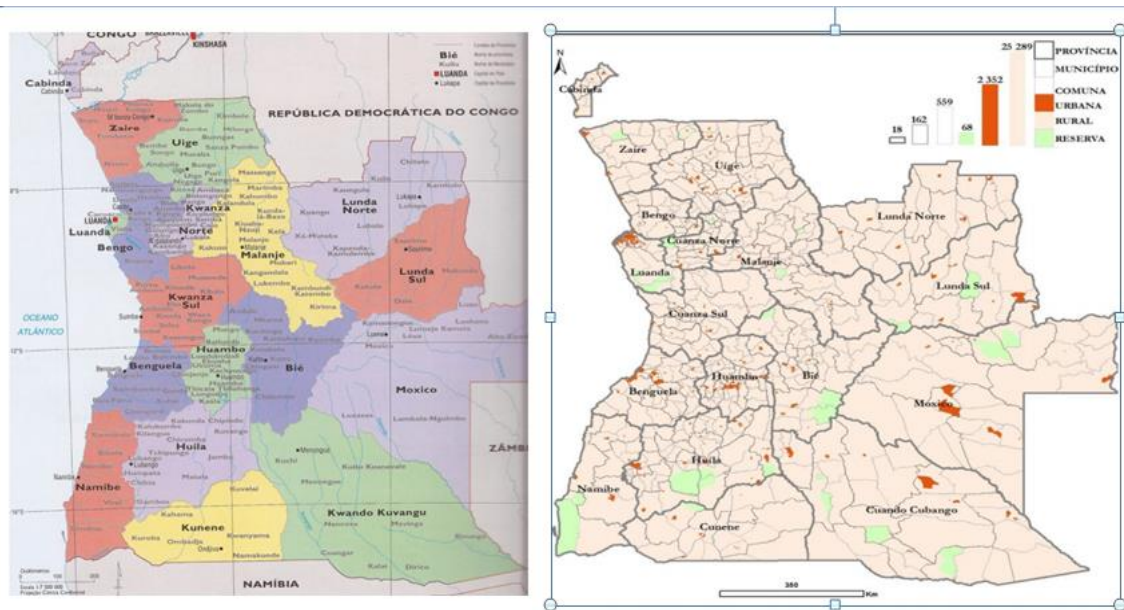
The standardization of the lexicon is called lexication (Kaplan, 1997; Nahir, 2002; Kaplan & Baldauf, 2003) or lexicalization (Bauer, 1983; Lipka, 1992a; Lipka, 2002a; Fernández-Domínguez,

2010). For Lipka (1992, p. 107) lexicalization is “the process by which complex lexemes tend to become a single unit with a specific content, through frequent use”. It is “a gradual, historical process, involving phonological and semantic changes and the loss of motivation” (Lipka 2002a, p. 113). Lexicalization can be classified as (1) “phonological” for example, the word “**A**rabic” (whose stress is in the “irregular” antepenultimate syllable) is expected be stressed in the penultimate syllable in the same way as “**i**talic” (mainly by second language learners); (2) “morphological” which explains the opaque relationship between “eat” and “edible” (as “edible”, a latin-based word, is strangely a synonym of “eatable”, an English-based word); (3) “semantic” describes phenomena whose relationship with their signifiers is neither natural nor logic, but arbitrarily comprehensible by use, e.g. in the relationship between “understand” (etymologically from Old English “to stand among or between”) and “comprehend” (etymologically from the latin “to seize with(into) the mind); (4) “syntactic” describes a lexicalization that derive from the transformation of two different grammatical class words by means of composition processes to form an entirely new compound word, e.g. “pickpocket” is not the act of picking a pocket, but either a person who steals objects in the other people’s pocket (or the act of doing it); (5) “mixed lexicalization” exists when one linguistic *item* , such as “lammas” (meaning harvest festival, St. Petre’s Day, a kind of loaf, etc.) may result in multiple lexicalizations at the same time (Lipka, Handl & Falkner, 2006, p. 6).

In fact, lexication involves the categorization of words into functional classes, deciding which words and idioms will be acceptable for formal setting and which ones will not, this list includes the knowledge about what words and expressions are taboo or swearwords. Lexication also involves the construction of specialized dictionaries to respond to the demands of science and technological development.

All the above discussion on lexicalization/lexication is important for this study as many examples of lexication occur with all human languages, Umbundu included. For example, the lexicalization of the following Umbundu words: “*ondisionaliu*” (from the Portuguese “*dicionário*”); “*putu*” (from the Portuguese “*Português*”) for “dictionary” and “Portuguese” respectively (Daniel, 2010, p. 13); “*palata*” (from the Portuguese “*prata*”) and “*ngalasa*” (from the Portuguese “*graça*”) for “silver” and “grace” respectively (Inverno, 2009, p. 128) are examples of morph-phonological lexicalizations as a result of the process of indigenization or nativization that Portuguese loans have undergone with time (cf. Anwar, Rasool & Kamran, 2020).

Figure 4: Two maps showing discrepancies in the orthographical representations of local toponymy.



Note. Two ways of writing the names of toponyms in two maps of Angola published in the same year, illustrate the still chaotic State of the orthography of native languages (Costa et al., 2016, p.11; INE, 2016, p. 28, respectively)

Figure 4 presents two maps of Angola in works that were published in the same year (2016). Both maps show the same toponymic references, but the way they are written differs. For example, “Kunene” and “Kwando Kuvangu” on the left hand map appear as “Cunene” and “Cuando Cubango” on the right hand one. As it is, the different orthographical approaches to the writing of toponymic entities evidentiate the lack of a consensual agreement on which standard to follow, which, in its turn, makes it difficult to write the local languages with a learner-friendly/teacher-friendly/user-friendly perspective.

Furthermore, the option for a radical move to a newly-created system – mainly when dealt with by specialists who belong to the language community may result in a more debated, consensual, informed, coherent sound-to-letter representation (Liddicoat, *loc. cit.*). So, two important criteria will be discussed in the following subsection as they are unanimously taken by scholars as orienting decisions on graphization. The first is the sociolinguistic one whereas the second is the psycholinguistic one.

2.1.2.8. Two crucial criteria guiding graphization

The evaluation of the effectiveness of graphization must obey two criteria: the sociolinguistic and the psycholinguistic (Cooper, 1989). As far as Liddicoat (2005) is concerned, the sociolinguistic criterion establishes that the effectiveness of a script will be measured by the way the political, economical, ethnical, religious, status, acceptance factors, etc. are in harmony with the employed graphization approach and process, within a given ecology of the language under study, considering the material conditions available for such a complex endeavour. On the other hand, however, the psychological criterion considers that for a language script to be effective it has to facilitate reading, writing and it has to be a helpful instrument of skill transfer. Apart from that, it has to be technology-friendly, adaptable to the demands of the rapidly changing planet (cf. Cooper, 1989).

There are contexts in which determining a standard writing system involves the capacity to choose among a plethora of concurrent non-officialized writing systems. That is the case of *Soomaali* whose standard script came to be a politically imposed Latin-based one, despite the existence of the widely used and accepted *Far Soomaali* - a syllabary script – and an Arabic script. The choice of the Latin-based orthography depended on the belief, by a group of influential politicians, in the idea that the script is linked to modern, secularist and tradition-free world (Andrzewski, 1983). This explains the sociolinguistic criterion discussed above.

The psychological criterion, explains the plausible reasons that have been behind the improvement of chosen standard scripts. For example, according to Unger (1996) in Japan the reform was carried out with the purpose of diminishing the Chinese script called *Kanji*, while in China the reform was done for the sake of simplifying the quantity of strokes in the structures of the characters. This examples show that considerations of various factors are due to the fact that all the reasons that can be invoked to adopt/improve a given standard script should take into account that people may face difficulties to unlearn the systems of writing they already know to assimilate the newly proposed representations.

When the script system has been decided, a very important task should be taken into account: the necessary regular revision for updating. In fact, Liddicoat (2005, p. 998) explains that most times “the spelling conventions that exist in the language are felt to be too complex or too anomalous, or even because they no longer accurately reflect current pronunciations”. An example given has to do with the attempts Webster made to give English a more functional, realistic and easy to read writing system (Ayto, 1983). To some extent, a few Webster’s proposals were considered, but others were massively discarded. Besides, there is an example of the revision and exerting changes

in the French writing system. This attempt has failed completely, for the most influential users have understood that the way the French orthography stands with its complexities gives it a special characteristic of strength, making it only accessible to intelligent people (Schiffman, 1996).

In the above line of thought, in Angola, the Umbundu language writing systems have been massively determined by religious motivations (Diarra, 2003). On the one hand, Protestants (and most non-Catholic writing system users) write the sound [tʃ] using the letter <c> and this pattern has been adopted by the “Institute of National Languages” – a State related body (Diarra, 2003, p. 343). On the other hand, the Catholic Church members have adopted the <tch> for the sound [tʃ]. In that case, the same Bible has been written in two orthographic standards for the Umbundu language. Unless specialists in the area of language sciences are really involved and taken into account to study and propose ways of reaching an orthographical agreement of the language, this language may be very difficult to be taught and learnt at the government schools. The criterion that both Catholic and the Protestant orthographical systems take is more sociolinguistic (each group defending its standard) than psycholinguistic (justified with the functionality and the reader-friendliness on the part of the agents of the learning and teaching process), although each support group may also claim a psychological criterion on the Umbundu orthography.

Having shortly discussed the standardization of corpus focusing on the process of codification, it is now opportunity to look at another important aspect of *corpus* planning: *corpus* elaboration. With it, an incursion around the concepts of lexical modernization, stylistic modernization and renovation will be discussed.

2.1.3. *Corpus* elaboration

Corpus elaboration has to do with the task of giving the *corpus* the quality that is needed to not only make the *corpus* of the language robust, but also justify its preconceived qualification to respond to the demands of the moment insofar as the modern world is in a very rapid development in all the areas of life, including the area of Applied Linguistics and that of Communications (Cooper, 1989). *Corpus* elaboration is an attempt to standardize language towards modernization in style, modernization in the lexis and renovation (i.e. purifying the language, reforming the language, making the style simpler and unifying the terminology) (Gonzales, 2002).

In its turn, Lexical modernization is the process of language expansion. It encompasses the modernization of the corpus so as to tune both the newly-developed standard language and/or the ones which already exist. As a matter of fact, there has been a tendency to develop standard

languages so they can be translatable into more influential languages of the world, such as English, French and Spanish (cf. Liddicoat, 2005; Mühlhäusler, 2000; Dash, 2019). Seen from this perspective, lexical modernization processes should consider a meticulous selection of specialists with demonstrated experience and competence in the area of linguistics, in general, and language planning in particular so as to assist other scientific areas, interdisciplinarily, in the establishment of advised use of the technical terminology necessary for academic and public communication.

There are two methodological decisions to make in the development of new terminologies: either by appropriating a term among the ones already existing or by creating a new term (Liddicoat, 2005). So, it is always important to be cautioned for the possibilities of unnecessary linguistic imperialism (Phillipson, 1992) over the old standard lexicon. Often times, borrowings get nativized (see examples of lexication in the antecedent section) by influence of the host language. This process may represent the talent that the host language users have to adopt and yet modify the donor language to respond to the way that they articulate words.

Stylistic modernization is a habit-dependent process which results from the regularity of patterns of textual elaboration, mainly those textual productions proposed, accepted, used and primarily disseminated within an elite (usually academic group) and generalized for use (most often, institutionally lawfully-enforced) in the social transactions of all formal kinds (cf. Kaplan & Baldauf, Jr, 2012); Liddicoat, 2005). Stylistic modernization is essentially a top-down process. As such, stylistic modernization is a process, in language planning, which undergoes regular updating and/or renovations to conform to the dynamic world of science and technology. In the context of Umbundu, it can be said that Malumbu (2007) presents a more Catholic-based orthographic style whereas Daniel (2010) presents a more Protestant-based orthographic style.

Stylistics simplification is an enterprise that demands specific thinking skills from the language planners. In an article entitled “Translating Children’s Literature: Some Insights from Corpus Stylistics”, Dr. Anna Čermáková States that Corpus stylistics is the use of corpus linguistics techniques to the examination of literary [and/or factual, *emphasis mine*] works by means of “[...] comparisons – the text under study is being compared to another text and often to a large collection of texts [...] called reference *corpora*” (Čermáková, 2018, pp. 121 – 122). For example, the writing style in a book of Law can be compared to that of another book either within the area of Law or compared to another of the Aviation speciality. A clear example of stylistic simplification can be understood from the process of development of a planned “artificial” language – Esperanto – as compared to a process of “natural” language planning (Tonkin, 2015, p. 122). The process of

stylistic simplification should be carried out with the intention of making language use less formal, but not necessarily extravagant or populist. The vocabulary, the syntax, the phonology, the discourse, and other linguistic, semantic and pragmatic features should be ones which can be fairly easily attainable and available for the common reader to understand.

Renovation is another aspect of modernization. As Cooper (1989, p. 154) puts it, renovation is “an effort to change an already developed code, whether in the name of efficiency, aesthetics, or national or political ideology”. In the dynamics of knowledge construction, a need for renovation is crucial, and it leans on the need to purify writing systems, reform them, simplify their style, and unify their terminologies and, ultimately, the process may consider the internationalization of the standard language. Many attempts have been made to renovate aspects of the Umbundu language at the level of its orthography (see Jimbi & Sicala¹⁹: 2020; Malumbu²⁰, 2007, just to mention a few).

Very briefly, purification of the *corpus* can be understood as a renovative process whose aim is to make language clean of any “contamination” which might have come to the standard language *via* borrowings, coinages, calques and other foreignisms (Rogers, 2015, p. 111), under the discernment that a language can be, by some means, contamination-free (cf. Liddicoat, 2005, p. 1003). As purification can be subtly carried out through the use of censorship as method (i.e. control undesired words and prohibit their use in public). Another method is eradication (i.e. the literary production should be clean of any undesirable linguistic item. A third way of purification is the preventive action of making intruder words not to enter the standard lexicon. Finally, there is the replacement of “intruder” elements by recycling dead words, resorting to the dialects for capturing possible substitutes, or else create totally new words (cf. Liddicoat, 2005, p. 1003 ff).

A very important subtopic in *corpus* planning is that of reform. Nahir (1984, p. 113) has defined language reform as “the deliberate manipulation of language [...] through a simplification of orthography, spelling or lexicon, as well as to serve the underlying political, socio-economical, cultural and ideological tendencies of the community.” A very practical example of recent reforms occurred in China because of a felt need to establish a more modern, and practical writing system for the people, adopting a latinized orthography – the alphabet (Dwyer, 2005). The reform had unconditional support from the most important political elite (DeFrancis, 2006).

¹⁹ Jimbi and Sicala (2020) propose an orthographic entrepreneurship towards the preservation of Umbundu. The idea is that by resorting to writing systems other than the Portuguese, consistent sound-to-letter codification can be guaranteed as there are sound articulations of the Umbundu language that the Portuguese alphabet cannot cover.

²⁰ Malumbu (2007) proposes a grammar in the Umbundu language which translates all the technical terminology of grammar into Umbundu, in a way that no one has ever attempted before. It can be said to be a most recent attempt towards the terminology development of the Umbundu grammar.

Terminology planning is an area of corpus planning related to the admission of terms into a language according to the area of use. According to Zarnikhi (2016, pp. 12 – 13),

“[T]erminology planning deals with terms and their related issues, [...] ranging from creating new terms to standardizing the existing ones, and to present them in the form of terminological products to the target users proportional to their sociolinguistic needs and aims, from stable linguistic situations to lesser-used languages.

From the quote above one would want to know what fundamental purpose terminology planning serves. In fact, “[t]he aim of terminology planning is (should be) the improvement of scientific and technical communication. In practice, terminology planning is also an element of *status* planning and common language politics”, (Hermans, 1991, p. 688).

Still, two more crucial concepts need to be presented here for enriching the current discussion, as they often occur in the processes of terminology construction: demotization and destandardization (Dovalil, 2020; Pillière *et. al.*, 2018; Willems, 2014). The former refers to the “revalorisation, ideological upgrading, of ‘low-status’ language to ‘best-language’” (Coupland and Kristiansen, 2011, p. 28). One important factor contributing to the elevation of non-standard varieties is the work of the *media* with their specialization in facilitating the dissemination of the elective standards throughout the different sectors of the society (which is not always an easy task, at all, if a conservative stratum of the society does not reach humbleness to collaborate towards implementation of ‘revolutionary’ views and proposal to ameliorate communications in the society at large), guaranteeing, in that way, the more or less advised, ‘slow-and-sure’ acceptance of the standard in promotion by new adepts from different strata of the society (*loc.cit.*).

The latter of the concepts above – destandardization – stands for the process whereby a prestigious standard language “loses its position as the one and only ‘best language’ [as identified by determined kinds of ‘well-bred’ users, *emphasis mine*]” for its being adopted and applied in non-traditional contexts, where it undergoes a process of “weakening [for having become of use by ‘strange, non-qualified’ users, *emphasis mine*], and an eventual abandonment of the ‘standard-ideology’ itself” (*loc.cit.*). The possible abandonment of ‘standard ideology’ happens for the sensation of “no more need” for differentiation between ‘standard’ and ‘non-standard’ varieties, as the standard language moves from the high classes’ exclusive consumption to the common people’s everyday utilization, and opportunely appropriated by the latter. As a matter of fact, in destandardization, the normative variation does not really come to weaken. It seems that it only becomes generalized in contexts it did not use to be expected. Consequently, language users, independent of which social class they usually belong to, will be equally using the same standard

variety, reaching a stage of language use equilibrium (cf. Coupland & Kristiansen, 2011, p. 29). This equilibrium may bring about the interference of linguistic players, items and practices which the preservers of the standard language are not prepared to acknowledge. Moreover, important part of the preservers (say teachers and media editors) may feel unmotivated to keep on correcting for the standard variety preservation (Dovalil, 2020), leading to the appearance of imminent standard variety, which has, in turn, the susceptibility to deteriorate the so-called normative [official] standard variety. The emergence of the Angolan Standard Portuguese discussed in Undolo (2013) and in Kandjambu (2020) is an example of destandardization of the European Portuguese standard in Angola.

To understand the process of destandardization and its implication in terminological unification within the corpus planning process, Vít Dovalil presents the following reflective question:

Who stops managing (= intervening in) whose language use, based on which expectations, in which situational contexts and social networks, why, with which intentions and consequences, when radical weakening resulting in the eventual abandonment of the standard variety is going on? (Dovalil, 2020, p. 192).

This question has its crucial role of leading to the consciousness of the natural emergencies of new forms of language manifestations leading to their natural acceptance as potential standard varieties which can eventually be officialized.

A successful terminological standardization process involves seven functions, namely “research, standardization, dissemination, implementation, evaluation, control and updating” (Cabré, 1999, p. 49). For these functions to be accomplished terminological standardization process undergoes eight stages in the following sequence:

- a. Analysis of the terminological needs of situation in accordance with the overall situation, and selection of the most suitable strategies for intervention
- b. Preparation of the terminological research plan adapted to the needs of the environment in question
- c. Preparation of the terminology with the participation of relevant users
- d. Standardization of the prepared terminology
- e. Choice of the most suitable format and presentation for the prepared terminology
- f. Implementation of the terminology in practice by suitable policies
- g. Monitoring the use of the terminology
- h. Constant updating of the terminology (Cabré, *loc. cit.*)

Not least important to remind is the fact that unification should attend to terminology, as well as the unification of the orthography, morphology and syntax (Hornberger & King, 1999). Terminology can be used at various different levels of knowledge production as observed, analyzed and advised by specialists.

One might well question the pertinence of such a discussion for a work that deals with the Umbundu language. The explanation is that any living language lends to the global knowledge's codified experiences in the areas of medicine, technology, warfare, and so forth. So, it is more than acceptable and expectable to explore the terminology that Umbundu may bring to the areas of science and technology with important inputs once it is well explored and put to use. Every language has a range of terminologies according to different areas of activity in the community, be day medicine, workmanship, education or the army. The work of Malumbu (2007) is an example of it. The introduction of important grammatical terminology in his Umbundu grammar renders a qualitative advance in terms of the insertion of Umbundu at the level of understanding that can pair with any language learner and/or researcher. Before Malumbu's grammar was published many of the terminology used were Portuguese ones as no one had thought that some technical and scientific terminology can be retrieved from own cultural *memoire*. To sum up, the importance of this discussion here rests in the fact that unification of terminology is a process that has to include the understanding of terminology planning as well as both the concepts of demotisation and that of destandardization. The two last concepts are crucial elements determining the emergency of new possibilities to envisage standard languages. New terms result out of the awareness of new language practices in the linguistic communities. Depending on how they are promoted, they may most probably enter the new lexicon and be unified to appertain the vast repertoire of a standard language.

Internationalization is a complex and multifaceted concept which englobes various areas of specialization and domain in the productively globalized world (see Hénard, Diamond & Roseveare, 2012; Doiz, Lasagabaster & Sierra, 2011; just to mention a few). In terms of communication of experiences around the world, internationalization has become a glottophagic process threatening the existence of many already endangered "minority" languages, with English, French, Spanish and other western languages as the principal instruments of acculturation and calculated establishment of a "neoliberal" agenda, and the consequences deriving from this posture (cf. Doiz, Lasagabaster & Sierra, 2011, p. 346; Clark, Haque & Lamoureux, 2013, p. 7 ff). In Language corpus planning, internationalization has been defined as a "process of improving computers, systems and internet

protocols to accommodate the diverse language needs of the world” providing for the “localisation and computing” in the world languages (Osborn, 2010, p. 49).

Internationalization has been used to facilitate and guarantee technological atmosphere, including the insertion of ICT (information and Communications Technology) adapted to localization and plurilingual tasks organization and orientation, technical standardization (for example, the adoption, use and development of the Unicode – a universal set of characters in computer hard disks) to facilitate the creation of a writing system which can respond to all possible world languages’ technical needs. So, internationalization should capacitate the intellectual community to “edit in diverse languages and scripts” for the world awareness of new linguistic events (Osborn, 2010, p. 11).

As to the African linguistic context – mainly in the multilingual context of Umbundu orthography – many limitations have been debated insofar as the use of Unicode software is concerned and this naturally affects Umbundu as a disadvantaged language in terms of writing responses from computers if compared to Portuguese, the official language in Angola. In fact, Osborn (2010) lists the following problems:

1. Both computer users in general and ICT engineers in specific do not demonstrate a full awareness of and knack in the use of the Unicode for the edition of African language texts.
2. Computer keyboards are not fully responsive to the immediate needs of the African native languages,
3. It is difficult for the Unicode to accurately represent oral language in writing towards preservation of the semantic and pragmatic value of the cultural aspects embodied in the African languages. Principal issue is to do with the diacritics as used within Latin-dependent script system.
4. Available computers and other gadgets may not be adapted/adaptable, in terms of disk size and keypad, to accommodating the amount of new Unicode information necessary for updated edition of local language texts.

According to Osborn (2010, p. 53) there are perceived “needs of African languages [...] to be, fully addressed”. Furthermore, a consideration is put on the strategic use of various technological standards to make both localization and internationalization come true for all possible living languages. A possibility to record language in the written form and/or saving a variety of recorded unities can bring hope for those languages which are in danger and facilitate their restoration. In fact,

such a kind of experiences has been led by Christian Missionaries in Nigeria whose goal is, among others, that of

“[s]aving endangered languages and empowering young speakers in order to make knowledge more easily accessible to them. Writing, establishing orthographies and publishing primers and storybooks have become some of the preferred strategies of minority communities to put their languages and linguistic heritage on the map, make themselves visible and participate in local as well as global (via the internet) multilingual discourses (Storch, 2013, p. 128).

An example of internationalization is the use of Umbundu in songs in Brasil by brasilian singers as with the song “*Imba*”²¹ which means “sing”, or another song “*Ku yesu wange*”²² which means “*To my Jesus*” sung by a South African choir. The use of the internet to upload songs in Umbundu and respective lyrics makes it easy to internationalize the language and its orthography, as many Angolan migrants can find and share them with others in the countries where they are. The internationalization of Umbundu also happens within the context of the Christian mission as masses hymns are generally sung in the native language and taken by missionaries to various other parts of the globe. Moreover, research work and publications about the Umbundu language may be found interested on the part of linguists around the world, thanks to the development of the technological means available nowadays.

To sum up, internationalization of corpus planning is a very important aspect of language planning and it often goes hand in hand with a detailed study of orthography. As the written form of most of the Bantu languages is orthographic, dependent on Latin-based typeface, an effort will be made, in the sections and subsections that follow, towards exploring the possibilities it offers for the amelioration of the orthographic representation of African languages.

The chapter on *corpus* planning within the language policy process has discussed necessary aspects of language planning. It has been necessary to briefly discuss the concept of both language planning and language policy so as to clarify the confusion these two terms may cause as a result of the misleading development of these concepts in the literature. Also, it was necessary to revisit the types of language policy that there are: the endoglossic and the exoglossic ones. Another aspect that needed attention was the description of language planning with its components such as graphization, lexication, internationalization, the nature of terminological decision, *inter alia*. The next chapter, however, will be devoted to the discussion of orthography as a component of language planning

²¹ Available at <https://www.youtube.com/watch?v=fWTjJoRzbBw>

²² Available at <https://www.youtube.com/watch?v=hNzcRWmgMb8>

(graphization). The chapter will, in a more focused way, discuss the concepts of shallow and the deep orthography as well as the criteria for an optimal orthography design process.

3. CHAPTER THREE: ORTHOGRAPHY: NATURE, TYPOLOGIES AND PROCESSES

The previous chapter was dedicated to the discussion of language planning, language policy and the way the processes of corpus planning evolved from earlier studies. This chapter will offer a discussion of the aspects of orthography as a vast area of study involving a historical background and experiences of various kinds.

The term orthography refers, to human mechanisms of printed symbols also called *scripts* or *writing systems* ideologically supported (Yifan, 2019, pp. 91 - 93), used to structure linguistic items to communicate ideas in social interaction. It has been conceived as, primarily, “correct or accepted writing and spelling”, and it also refers to the research focused on “letters [and other printed symbols] and how they are used to express sounds [or a sequence of sounds]... whether considered 'true' and 'correct' or not” (McArthur, 1992, pp. 732; Murphy, 1996, p. 46)). In fact, the term “orthography” has evolved from the Greek *orthós* which means “correct” and *gráphein* which means “to write” and/or spell (Bussmann, 2006, p. 845; cf. Algeo, 2010, p. 36). Seen from this perspective, orthographies have been conceived to be accepted systems, and the study of orthography is (or at least should be) made in a way that each letter (and/or other written symbol) corresponds to segment of sounds, so as to facilitate readers to apprehend meaning from the texts they read.

3.1. The cenemic and the pleremic orthographies

Human writing systems can be divided into two main groups. On the one hand, there is the cenemic system in which scripts are related to phonic units. This category can be subdivided into five types (see Baroni, 2011, p. 128; Karan, 2006, pp. 34 - 61): the abjad, the syllabary, the alphabet, the featural and abugida. On the other hand, there is the pleremic system. Baroni (2011: *loc. cit.*) offers three such types of pleremic writing: pictograms found in the Egyptian literacy; ideograms – Chinese scripts and the logographic systems are morphosyllabaries which use one character, such as “☀” representing the sun (Daniels 2001, p. 43; Karan, 2006, pp. 34 - 61).

Orthography may be consciously or unconsciously used both as a means of deviation and a means of preservation, given the fact that the Roman scripts that pervade the vast majority of world’s language alphabets to date cannot be sufficiently responsive to the phonological and phonetic

demands of most of the languages which have, so far, been encoded or in the process of codification, particularly the non-European languages (Roberts, 2014; Miethaner, 2000; Dalby, 2004). Additionally, as in the Comoros Islands, the orthography of languages has undergone processes influenced by “identity, politics and linguistic ideology” (Ottenheimer, 2013, p. 2), because different colonizers wrote the archipelagoes’ languages according to the linguistic norms used in the countries they came from, namely the norms of “Swahili, Arabic, Hindi, Malagasy, Portuguese, English and French”, depending on which colonizer has controlled which of the archipelago’s part (Ottenheimer, 2013, p. 1), resulting in different ways of writing the languages of that island.

Be they cenemic or pleremic, orthographies can be designed to facilitate the direct, one-to-one relationship between scripts and sound segments or making it difficult for this access to operate during the reading process. This is a discussion to be undertaken in the subsection that follows.

3.2. Shallow versus opaque orthography

Text writing and reading in different languages has demonstrated that, more often, the relationship phoneme – grapheme does not confirm a one-to-one relationship (Murphy, 1996) and the orthography of the alphabetical systems is often very contradictory when it comes to fulfill their role of making the principle of one-letter-to-one-sound, in producing transcribed orality, observable (Jaffe & Walton, 2000; Bussmann, 2006). In this regard, the literature on orthography has shown that, as far as consistency of a writing system is concerned, two terms must be taken into attention: “raw phonemicity” and “adjusted phonemicity” Bontrager (2015, p. 27). The former has to do with how one phoneme is expected to relate to one grapheme only, while the latter explains the reasons for the irregularities that may happen in relation to an expectable, frustrating one-to-one relationship between the graphs and the sounds, as influenced by the co-text and the context. For example, the character <ç> in Portuguese has one consistent phonemic value – [s] representing a case of raw phonemicity, whereas the characters <c> and <g>, each one of them represents more than one phonemic realization in both Portuguese and Spanish – sustaining the concept of adjusted phonemicity (cf. Bontrager, 2015, *loc.cit.* for more detail). In fact, (<c> can be realized as [k] if followed by vowels X & Y and as [s] if followed by vowels [i] and [e]. The Portuguese word “comece” (“start”) is pronounced as [kɔ 'mese], with the same letter <c> pronounced as [k] before the letter <o> and pronounced as [s] between the letters <e>. Similarly, the word “geografia” which means “geography” has got two letters <g>, with differing pronunciation on the base of which position they take in the written word. Before <e> it sounds [ʒ], while before “r” it changes into [g].

When it comes to the Umbundu orthographies, for example, based on the Portuguese alphabet, the official language of Angola, both the character <c> and the combination <tch> assume the phonetic value of [tʃ] respectively in the Protestant version and Catholic one, so that one can write “cove” [tʃove] and “tchove” [tʃove] to mean “yours”. As to the letter <g>, it takes the sole phonetic value of [g] throughout in both Protestant and Catholic orthographical systems (MPLA, 1980, pp. 101 – 105; Jimbi, 2018, p. 902; Daniel, 2010; Valente, 1964, just to mention a few).

The example with the letter <c> shows that the Umbundu orthography is opaque in both Protestant and Catholic versions; (Schadeberg, 1986, p. 430 ff). Its opacity also resides in the fact that a learner’s/user’s only language of instruction at school and most of the official information is in Portuguese, which makes it difficult for users to read or write in Umbundu as they already have to face the opacity demonstrated in the Portuguese writing system which heavily interferes with the reading and/or writing process of Umbundu. As with the use of the letter <g> with a consistent value of [g] in both the Catholic and the Protestant orthographies, the problem consists in the fact that students learn Portuguese from the primary school which may lead them to an overgeneralization of the [ʒ] that happens when <g> is positioned before <e> or <i> (Prista, 2003, pp. 207 – 208). Consequently, a learner of Umbundu may have problem pronouncing the word Umbundu word “omange” [ɔ ‘mʌŋge] (MPLA, 1980, p. 105) and pronounce it as [ɔ ‘mʌŋʒe], a non-existing word in Umbundu as an overgeneralization from Portuguese words such as “longe”, “viagem” (far and voyage respectively). On the other hand, the writer has also to decide on how to write the words, so that in terms of the graphemes to be produced raw graphemicity and adjusted graphemicity correspond to raw phonemicity and adjustable phonemicity respectively. It is, by now, important to State that phonemicity and graphemicity undergo different processes in that, for example,

[A] hypothetical code in which <c> and <K> may both be used to stand for /k/ [as illustrated above]. Any time a reader sees either letter [in the *feedforward* direction], he/she will readily know to pronounce it /k/, so the ambivalence is no detraction in phonemicity. Anytime a writer wants to spell a word containing /k/[in the *feedback* direction], on the other hand, he/she may not necessarily know (at least not based on purely phonological criteria) whether to use <c> or <k> , and so this does constitute a detraction in graphemicity (Bontrager, 2015, p. 28; Protopapas & Vlahou, 2009, p. 991).

It is, in addition to the above, that the concepts of shallow and opaque orthographical systems can be discussed. It is a shallow or transparent orthographical system, the one of which each letter of the alphabet will match with one sound segment as with Italian, Portuguese, Greek and Spanish, whereas an opaque or deep orthographical system will be determined by the multifaceted representation of one letter for more than one sound in the same language or one segment of sound represented by more than one letter, as dictated by the context in which a graph occurs as

experienced with English, Danish, and French orthographies (Frost, 2005; Protopapas & Vlahou, 2009; Elbro, 2005; Goswami, 2010; Miller, Kargin & Guldenoglu, 2012). In the case of Umbundu, the opacity occurs in the two orthographical versions, which makes it difficult for students to discriminate which version to follow, let alone the fact that there seem to have been no consensus about which orthography to adopt for the education system (Jimbi, 2018).

In the development of orthographies for languages – both for the first time and for the subsequent phases - there have been orthographic criteria used by linguists and other language planners. The following subsection will briefly look at some important criteria for devising optimum writing systems.

3.3. First proposed criteria for optimum writing systems: A brief account

Basically, the majority of known languages of the world have been written on the basis of the Graeco – Latin alphabet (Lüpke, 2011). Ideally, as discussed above, graphemes of this alphabet were invented to represent the sounds of languages in a one-to-one correspondence to facilitate rapid reading and access to knowledge (through language). However, the historical development of the existing orthographies for the world idioms have been adapted and influenced by the social, political, historical, economical, cultural, religious and intellectual contexts prevailing in the time of their adoption (Lüpke, 2011; Oko, 2018; Alpatov, 2017). This State of affairs made it possible for researchers in (socio)linguistics to discover and agree on principles that inform orthography promoters when devising written forms of languages, be they new or revised orthographies which, again, should not be taken as being applicable to all languages, as each language is a unique one (Malone, 2004).

A point to consider is that three are known as the first proposed criteria underlying the design or adoption of the orthography of a given language: the phonetic criterion, the etymological criterion and the historical criterion (as discussed in Teixeira, 2014, p. 55; Sgall, 1987, pp 2, 21). The first criterion postulates that each sound segment of an utterance should correspond to one alphabet's character (shallow orthography). It has even been Stated that Leonard Bloomfield, a well established American linguist, defended this principle and considered those who did not embark on adhering this principle to be “ignorant” and, proponents of this principle consider that “the road to perdition is paved with irregular letter–sound correspondences” (Venezky, 2004, p. 141; Elbro, 2005, p. 34). In

fact, most African languages were written on the bases of the traditions that were used for the colonizers' language alphabet (Batibo 2000, p. 153).

In the same way, Umbundu was written, as much as possible, according to the orthography of Portuguese, although some adaptations had to be made given the lack of letters to fill for some specific Umbundu sounds that the Portuguese language does not have. For example, the use of <k> for <c> has been adopted in both the Catholic and the Protestant orthographies to represent the sound [k]. In that case "Caluquembe", a toponymic element of Angola (Gonçalves, 1960, p. 72) should be written "Kalukembe" (Cf. Chimbinda, 2015; Gomes, 2016; Pedro et al., 2013; IECA, 2009; Malumbu, 2007). This effort has been made in order to make the orthography shallower, and so fulfilling with the phonetic criteria.

As Angolan native language teaching and use had been lawfully prohibited by the Angola Province's High Commissioner José Mendes Ribeiro Norton de Matos's famous Decree 77, 1st Series, No. 50 in 1921 (Matos, 1921, p. 43ff), the Umbundu orthography, most probably, came to be influenced by the orthography that missionaries adopted earlier for other African languages in other colonial territories under France and England control, where African languages had already been coded (Hambly, 1934, pp. 234 - 252) than resorting to the Portuguese orthography. Protestants who became totally disbelieved in the eyes of the Portuguese administration, by the year 1940, for their focus on the development local languages, culture and nationalist sentiment were the first victims of the "Missionary Accord" (Neves, 2007, p. 516). In fact, it has been stated that the first printed material in Umbundu is called "Vocabulary of the Umbundu Language: Comprising Umbundu-English and English-Umbundu; Lists of Three Thousand Words Used by the Inhabitants of Bailundu ... of West Central Africa" and it was published in Boston, Massachusetts, in 1885, by William Henry Sanders and William Edwards Fay, both American missionaries (Neves, 2007, p. 513 ff).

Anyhow, the one-to-one correspondence principle, however simple and economical it appears (Webster, 1967, 1789; Pitman, 1905; Barnitz, 1978; Venezky, 2004), has not shown to be adhered by practitioners and developers of writing systems. According to Venezky (2004, p. 142) the lack of firm centralization of the control over the spelling rules and the absence of special bodies linked to spelling-related issues which facilitated "scribes, orthoepists, lexicographers and the like" to determine the various ways of spelling. In France, on the other hand, the *Académie Française* – made of linguists and other language-related agencies whose intention was to ameliorate and give the French orthography the regularity it deserved - ended up scrutinized by the strongly influent,

conservative political sector of the epoch, mainly the Ministry of Education, to impede any possible changes in the writing system. These practices have been transported to the colonies these countries had around the globe and influenced the way many non-European languages' orthographies were devised and still are. Furthermore, African languages' orthographies have been designed between simplification and complexities (c.f. Karan, 2006) depending on which tradition the proponents of the African languages' alphabets have considered (cf. Batibo, 2000).

The etymological criterion (often confused with the historical one) can be conceived as “[t]he choice of one out of different letters for similar sounds [...] based on the derivation or older form which was in use when sounds that are similar now, could still be distinguished clearly” (Venezky, 2004, p. 148). In fact, the word etymology (from the Greek *etymon*, meaning “true sense/original meaning”) from which the word “etymological” originates has been defined as the understanding of “what steps does such and such a word come to have the meaning in which it is actually found, what is the earliest source to which it can be traced, and what are the cognate forms either in our own or in related languages” (Wedgwood, 1859, 2009, p. i). More than taken diachronically, etymological criterion, here, is to be taken synchronically, so as to avoid potential confusion with the historical criteria below. For example, the word *technique* has retained the sound [k] in the sequence “ch” from Greek and then Latin (i.e. *techniká, technicus, respectively*), having come to be naturalized more or less as they are in French. This principle results from the considerations of the origins of the orthographic arrangements of the sounds of the language, i.e. the due respect given to the imported characteristics, mainly when features currently used in one language can be found inserted in an earliest source language. This feature includes prefixes, suffixes, or entire words and expressions. Similarly, for the Umbundu language the etymological criterion is more evident in proper names as codified and imposed to be official in identification cards and other official documents created before independence. For example, in Gonçalves (1960, p. 52), the toponymic elements “Bailundo” and “Bié” represent a corruption of “Mbalundu” and “Viye”, respectively, as explained in Gomes (2016, p. 35). It can be stated that the official way of writing the word in the colonial period has remained up to now in charts and the administration panels and labels.

The third criterion – the historical one – is defined by the fact that the way the word is written can be explained by the linguistic community *habitus* (Santos, 2018, p. 261) with the language in a timeline. In this case the explanation about the orthography of a certain word cannot only be done on the basis of the phonetic or the etymological principles. Rather, it can be explained by considering its diachronic development within the linguistic community. This criterion helps people understand (a)

how orthography “got to be that way” in the course of time; (b) “that many of the irregularities in today’s language [orthography, here] are the remnants of earlier, quite regular patterns” which explains the famous joke linked to the English orthography that “fish might be spelled *ghoti* (*gh* as in *enough*, *o* as in *women*, and *ti* as in *nation*)” (Algeo, 2010, p. 17). Unlike the etymological criterion which concentrates on where and how a certain orthographic characteristic originated, this one focus on the changes that certain orthography has experienced with time.

The historical criterion with the Umbundu orthography is vast in examples. In a PhD thesis, Teresa dos Santos made a clarifying list of how some words passed from one aspect to another as the colonial administration wished (dos Santos, 2015). For example, The Umbund words “Katavola” has become “catabola”; “Lúvia”[*sic.*] has become “Lúbia” and “Viye” has become “Bié” (dos Santos, 2015, p. 109). The advantage of this criteria is that it can help understand what may have influenced the change of the [v] to [b], so clear that in the timeline, the natural difficulties of the settlers to make it right in their own language’s articulatory system, has made them “force” the change from “*Viye*” to “*Bié*” and from “*Katavola*” to “*Catabola*”; “*Luvango*” to “*Lubango*” as a sign of both submission condition and ability to be civilized. Another study carried out by Brissos (2011, pp. 105 – 111) may give a hint in that most of the Portuguese settlers of the Centre and southern part of Angola may have come from the northern and the South East part of Portugal where betacism (“*betacismo*” in Portuguese) – the inexistence of the articulation of the phoneme [v] which becomes substituted by the articulation of the phoneme [b] – making speakers to articulate [b] mainly in words carrying <v> like “*Viye*”; “*Kuvale*” and “*Luvango*”.

All in all, the understanding of the aforementioned criteria is crucial and applicable to the study of the Umbundu orthography. It stands evident that the phonetic criteria directs to a simpler and more economical way of writing, although many factors may influence the way orthography for a given language is tackled by different agents in and outside the government sphere. The etymological criteria help the specialists to understand what may have motivated the existence of a way of writing and how, where and when originated. The historical criteria is advantageous in that it may help researchers to understand what, in the timeline, is behind the appearance of a certain orthographic feature of a language. All this gives orthography an interdisciplinary and complex area to deal with.

The following subsection, as it will be seen, is very similar to the previous one. However, it concentrates on the demonstration of the principles underlying the devising of orthography as a practice in Africa in general and in Angola (for Umbundu), in particular.

3.4. Principles guiding orthographical systems in the African context

Apart from the above principles of orthography design, other ones have also been revisited by different researchers more recently. Although these principles tend to overlap and complement each other, they can be summarized as follows: (a) principle of high level of motivation; (b) principle of high level of representation; (c) principle of high level of learnability; (d) principle of high level of transferability and (e) principle of high level of replicability (see Karan, 2006, p. 64 for additional explication of each type of maximum).

A priori, Gary Simons had proposed the “[p]rinciples of multidialectal orthography design”, under a number of maxima, namely: (a) the maximum of public acceptability (b) the maximum of “psycholinguistic acceptability” (c) the maximum of reduced confusion (d) the maximum of “simplicity” (e) the maximum of encounter of distorted systems (f) the maximum of phonemic distinction and dialects unmarkedness (g) the maximum of generalized ease (see Simons, 1994, p. 325 for additional information on these principles).

Most recently, 6 principles of orthographical design have been considered in the African context. In fact, Longtau (2014, pp. 7 - 18) offers the following hierarchy of criteria which are both overlapping and consistent with the aforementioned proposals:

a) Accuracy or rigor. This principle, often called the principle of economy (Batibo, 2000: 154) respects the beneficiary community’s awareness of the characteristics of the orthography to be proposed, as their contribution may guarantee the safeguard of the “maximum representation of speech” and the “simplicity of rules” (Schroeder, 2008, p. 4) presented above. Rigor does not need to mean the total coverage of a one-to-one correspondence between phonemes and characters, given a plethora of writing systems and the factors influencing their choices. Nevertheless, orthography has to play its role of representing the language in the written form, as faithfully as possible, as a result of a competent scientific research, instead of depending solely on partial intuitional impositions. Orthography must, thus, have sufficient number of alphabetical characters to respond to the phonetic demands of the language it is designed to, as seen from the users’ comprehension standpoint. This way, rigor as an orthographical principle will allow for “trustworthiness” and guarantee a credible, transferable and dependable writing system with positive consequences on the criterion of convenience below (Ryan, Coughlan & Cronin, 2009, p. 312).

In the case of Umbundu, writing a word such as “Ñala” in the Protestant version [IESA, 2009, p. 245; Daniel, 2002:457) and its Catholic counterpart “Ñgala” (Mathew, 2 verse 24; Malumbu, 2007, p. 42), let alone the proposed <ɲala> and <ng’ala> (MPLA, 1980, p. 105; Pedro *et*

al. 2013, p. 2, 14) to orthography [ŋǎlǎ] may lead to question the accuracy of the *tilde* accent on the <n> instead of accenting <g> which is the one velar sound needing to be nasalised by using the diacritic *tilde* on the <g> as already proposed in Valente (1964, p. 19) and not the palatal <n> whose nasalization results in the articulation of [ɲ] as in <Spaña> after the Spanish orthography, a writing system created long before the orthographical inventory of Umbundu. This is a case of lack of rigor or accuracy to a great extent and it also threatens stability as discussed in f) below. To achieve rigor or accuracy Valent's proposal would better respond to the transliteration of the way it is pronounced in Umbundu, resulting in "Nǎlǎ". Until this accuracy is achieved, a good grasp of Phonetics and Phonology is required. Another example is the use of the digraph <ny> to represent the sound [ɲ] as conventional in the Protestant version while the Catholic version brings the digraph <nh>, so that the Protestant "omwenyo" (IESA, 2009, p. 310) and Catholic "omwenho" (Chimbinda, 2015, p. 132). For the first case, the way the sound [ɲ] is represented by <ny> is influenced by the Swahili and other Anglophone countries adaptation of orthography for the Bantu languages (Schroeder, 2008, p. 33ff). For the second case, the influencing orthography is the Portuguese one. Schroeder (2008) offers a list of alternatives which explains very clearly how orthographical options for African languages have been made which can help think of infinite possibilities to repair and introduce new ways of representing sounds of African languages in a more approximate way. It seems that a more economical way should be resorting to <ñ> for the sound above as it is used by a widely known language, Spanish, and one needs no more than one letter to represent a sound in a one-phoneme-one grapheme relationship, so facilitating "the new reader [the experienced reader, *emphasis mine*] and the learners of the language" (Venezky, 2004, p. 139). In fact, in the Olunyaneka²³ alphabet, <ñ> is used for the sound [ɲ] in a syllabus written under the sponsorship of The Namibe²⁴ Provincial Direction of Education, Science and Technology and the NGO COSPE²⁵ (Co-operation for the Development of Emerging Countries) (Ponte, 2012).

b) Convenience. This principle means that the proposed orthography must facilitate reading on the language users'. For this facilitation to occur, one must think of the necessary efforts to regularly revise and adequate the orthography, as much as possible, for the perceived needs of the language as manifested by the users. This facilitation is today better offered by the considerable advances in information technology which make computer overcome insufficiencies that the old typewriter could

²³ *Olunyaneka, olunhaneka or oluñaneka* (depending on who writes) is a regional language of South West Angola.

²⁴ A South Western province of Angola

²⁵ An Italian NGO "Cooperazione Per Lo Sviluppo Dei Paesi Emergenti" sponsored by the European Union.

not do before. Computers, nowadays, offer a plethora of possibilities through the UNICODE (Küster, 2019, p. 22) and the International Phonetic Alphabet resources. This principle overlaps and complements Gary Simons's aforementioned principles (c), (d), (f) and (g). Convenience demands revision and constant updating. For example, the Portuguese orthography on which the Umbundu orthography (mainly the Catholic one) is based, obviously, does not have a graph for the sound [tʃ]. For this phoneme to be graphed in that version a trigraph <tch> had to be adopted to write "*watchita*" meaning "has given birth to her/him" (Chimbinda, 2009: 31). In their turn, the Protestants adopted the graph <c> to write "*Cosi*" meaning "everything" (IECA, 2009, p. 310). The Portuguese administration adopted the digraph <ch> to write "*Chilessó*" and "*Chissamba*" which are toponymic elements (Gonçalves, 1960, p. 52). Apart from the above versions, other alternatives were found in de Oliveira (2012: 931) who presents "*tjila*" (playground) with the digraph <tj>, <ty> and <tx> representing the phoneme [tʃ] (see Diarra, 2003, Jimbi & Sicala, 2020).

Most probably, if the internet had been invented to a wider use at the colonial time that the Umbundu orthography was being devised, the involved agents would have known of the Czechoslovakian <č> which has been created to be permanent for the phoneme [tʃ] and adapt it for the Umbundu language. It would certainly appear more convenient to have a simpler, easy-to-read and easy-to-learn. As a matter of fact, the *Olunyaneka* alphabet – another regional language in Angola – has adopted <č> for the [tʃ] sound (Ponte, 2012).

c) Conformity. Orthography harmonization depends on this principle in that the need to conform the writing system to the real phonetic and/or phonological demands of Bantu languages in the colonial period, for example, was biased by the immediate technical and feedback facilities which were available to support and facilitate the Mission of Christ as the decisive factor for orthography design, independent of whether they captured the phonological realization of the autochthonous language users (Batibo, 2000, p. 153; Longtau, 2014, p. 14). However, technological advances have demonstrated that this kind of deficit can be overcome nowadays and it has become easier to adjust sounds to letters of alphabets for almost every language.

The harmonization of the Umbundu orthography, although possible, it is not an easy task because of three main factors that can be advanced in this study: a) the protagonistic factor related to who has created the first version; b) the production factor related to who produces more; c) the political non-specialist factor related to who has got the political influence to determine which version to choose. In that case, many have been attempts to harmonize its orthography (see for example MPLA, 1980; Pedro *et al.* 2013). The traces left show that the research on the creation of the

alphabets in “national” languages had not been made according to the scientific rigor. Only 3 informants from Huambo province; “2 informants” from Bié province and only 1 informant from Benguela province (MPLA, 1980, p. 99). Besides, the debate on the Umbundu alphabet took place in Luanda under the scrutiny of members of the sub organization of MPLA (UNTA - Proletariats’ Union, OMA - the Party’s Women Organization and members of the Central Committee of the ruling party) (MPLA, 1980, p. 100). Some important results include the approval of <c> to fill for [tʃ], the sound [ɲ] is represented by <ny> (as in the Protestant versions) and the sound [ɲ] was written as <ɲ> (this one can neither be found in the Catholic version nor in the Protestant one in the present days). 33 years after the first widely acclaimed attempt, another joint publication headed by the South African CASAS (the CASAS’ Monograph no. 251) and the “Instituto de Línguas Nacionais” (Institute of National Languages) came up with “Harmonização Ortográfica das Línguas Bantu de Angola” (Orthographic Harmonization of the Bantu Languages of Angola). In this one, instead of a harmonization as such, what happened was an attempt to revise the 1980’s project and change the <ɲ> into the trigraph <ng’> as used in Swahili (Pedro *et al.* 2013, p. 12). Looking at the list of participants of the study and the way the criteria of selection of participants was made, there has been very little to faithfully represent the population of each of the studied language, let alone the study of the Umbundu orthography whose number of participants does not seem to be representative enough as most of the participants are not native speakers of Umbundu (Pedro *et al.* 2013, p. ix). Besides, the most important universities and Higher institutes in the regions of Bié, Huambo, Benguela and Huila, where Umbundu is spoken was not taken into a necessary account on the part of the Angolan group (see Pedro et al. 2013, pp. ix – x). There does not seem to have been priests and pastors involved in the study, either. This is related to the factor c) presented above. Now, there also is a case of who wrote the first book (our factor a)) and whose population and Umbundu publications consumers are the majority (our factor b).

d) Acceptability. The language users have to give a positive feedback and accept the new orthography after the adoption of the writing system has been completely explained. This principle is important in that the community may always know aspects of the language that the researcher may not be aware of (cf. Fishman, 2006, p. 25). This can be guaranteed by the scheduling of regular meetings with the communities so that they can get involved and so participate in the creation of the orthography for their community (i.e. participatory orthography design).

Umbundu orthography has been devised by people who could hardly speak it (and seems to be so to a great extent). In that case, its orthography cannot be accurately designed if people who speak

it are not contacted, heard, and their voice taken into account. In times that the government is trying to insert it into the education system, the need to revise the orthography that will be used for the education system where both Protestants and Catholic members will study.

e) Word division. This principle is linked to the “[m]aximum recognition/transmission of meaning” (Schroeder, 2008, p. 5). For this principle to be fulfilled, knowledge in lexicology, lexicography, grammar, phonetics and phonology must be had and put into practice. As advised in Van Dyken and Lojenga (1993, pp. 3 – 5), the most challenging part of the orthography design process, after the alphabet has been established is word arrangements. This implies that decision should be made of whether text writing will be approached disjunctively or conjunctively. (Batibo, 2000).

As for the Umbundu word division rules, the CASAS and “Instituto de Linguas Nacionais” have advanced with 30 general rules for the unified standard orthography (Pedro *et al.*, 2013, pp. xii – xiv) which can serve as a starting point to revisit and decide on the nature of the Umbundu language as well as decide on whether to approach the orthography conjunctively or disjunctively. One problem with the rules that are established by the above institutions is the lack of examples from the languages they try to describe. For example, rule number 2 States that “The predictable strengthening of a vowel does not need to be indicated” and in the same paragraph a seemingly contradictory sentence States that “... in the majority of cases the meaning of certain words with vocalic strengthening is perceptible from the sentence context” (Pedro *et al.* 2013, p. xii). The point here is that it is not clear which vowel strengthening is predictable from the context of the sentence. For example, the sentence “iika” vs “ika” (“it closes” vs “gets used to”) “Okulila” vs “Okuliila” (“to weep” vs “to come along alone”). If words like these are not made clear orthographically, the reader may face difficulties to discriminate their meaning. So, it is important that examples follow the rules to reinforce learning.

f) Stability. This last but not least principle, goes that it is necessary that the orthography should be coated with a stable nature in terms of rules. It is therefore necessary that the users endeavor in consistency when producing their texts so as to avoid discrepancies in the texting products as developed by different writers. For the criterion of stability to be safe, Gary Simons’s “maximum of “psycholinguistic acceptability” needs to be considered, as stability is controlled by the psychomotor processes involving memorization, conceptualization, articulation/spelling and self-monitoring/editing. Moreover, stability can be guaranteed by institutions which have chosen to pave a way for the Umbundu language studies. Umbundu is taught for teacher to be at ISCED/Benguela

and at Benguela Jean Piaget Polytechnic Institute (see presentation and treatment of data in chapters 4 and 5). What is not known as yet is whether both institutions use the same version of the language. If there has not been any convention yet, then a general convention of the Umbundu orthography is urgent, if Umbundu has to be inserted in the education system in Angola.

The principles above are refurbished with the essential strength to support the design of an orthographical project which can more aptly be responsive to the local languages of Africa, be it at the level of starting original orthographies for a language, be it at the level of revision, updating and harmonization of existing orthographies towards the current demands of the linguistic communities. In fact, Portuguese, another widely spoken Angolan language has undergone at least 5 important orthographic reforms (i.e. one in 1911, one in 1931 to 1932, one in 1943, one in 1973 and the one in 1990) (see Pereira, 2012).

3.5. Creating an orthography description: A practical sketch

Creating orthography for a language obeys a basic outline (Hosken, 2003, p. 3ff). Hoskens's outlines include the following considerations to take into account:

- a. The sociolinguistic aspect.

This feature comprehends the knowledge and registration of the number of speakers of the language. As to the number of speakers of the Umbundu language, the percentage is of 23 (INE, 2016, p. 51), although it is wise to consider another important study which shows that 37% of the population speaks it (Monteiro, 2014, p. 28). Additionally, an interesting recent study has involved 173 respondents, namely 17 parents, 21 primary school teachers and 135 university students (Pinto & Silva, 2022, p. 7). That study has revealed that only 23, 50% of the parents (i.e. 4 parents out of 17) teach some Umbundu expressions to their children; 4,7% of the teachers , corresponding to 1 element, can speak Umbundu with fluency; 5,5% of the university students (i.e 7 respondents out of 135) can speak Umbundu fluently (Pinto & Silva, 2022, p. 9). Besides, there is important statistical information to take into account with regard to the number of people living in Angola today. The UNWFP's²⁶ "Angola interim country strategic plan (2020 – 2022)" attributes an estimation of 31.8 million people.

The statistics above is crucial to measure the importance of designing and regularly update the Umbundu orthography, given that if the INE results of 23% of Umbundu speakers are taken into account, that will represent around 5.931.476 individuals; in case Monteiro's (2014) percentage of

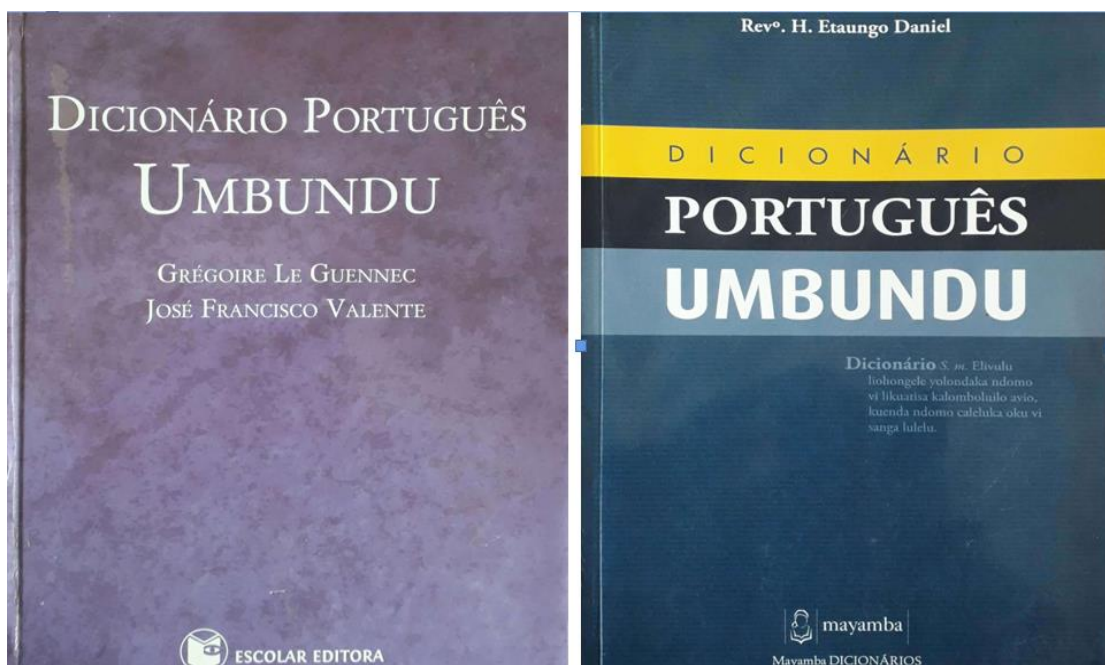
²⁶ United Nations' World Food Programme; https://executiveboard.wfp.org/document_download/WFP-0000108616

37% of speakers is taken into account, that will correspond to 9.541.939 people; with the estimation of the UNWFP's estimates the corresponding number of people who speak Umbundu will increase to about 7.314 to 11.766.000 people for the INE and Monteiro's results respectively. This number is more than enough to work over the orthography for this language.

Another important question to ask is "how many of the speakers have been literate in the language?" (Hosken, 2003, p. 3). By now, the percentage of speakers of the Umbundu language who are literate is difficult to calculate, due to the lack of official information and studies about it. Moreover, the statistics on language literacy in schools is exiguous and the interest to go about is amidst those who are for and those who are against (Sacalembe, 2021). As it can be inferred from Pinto and Silva (2022) the number of literate people as such is fuzzy.

Moreover, there is a preoccupation to know whether the orthography has been currently in use, by whom and to what extent. As far as use is concerned, Umbundu has been of more utility in churches which use them for the masses by means of the biblical registrations and the catechesis' book. Priest, pastors and the clergy members use them more frequently. Dictionaries and other materials have been written to help preserve the Umbundu language as can be seen in the samples below.

Figure 5: Dictionaries (Portuguese – Umbundu)

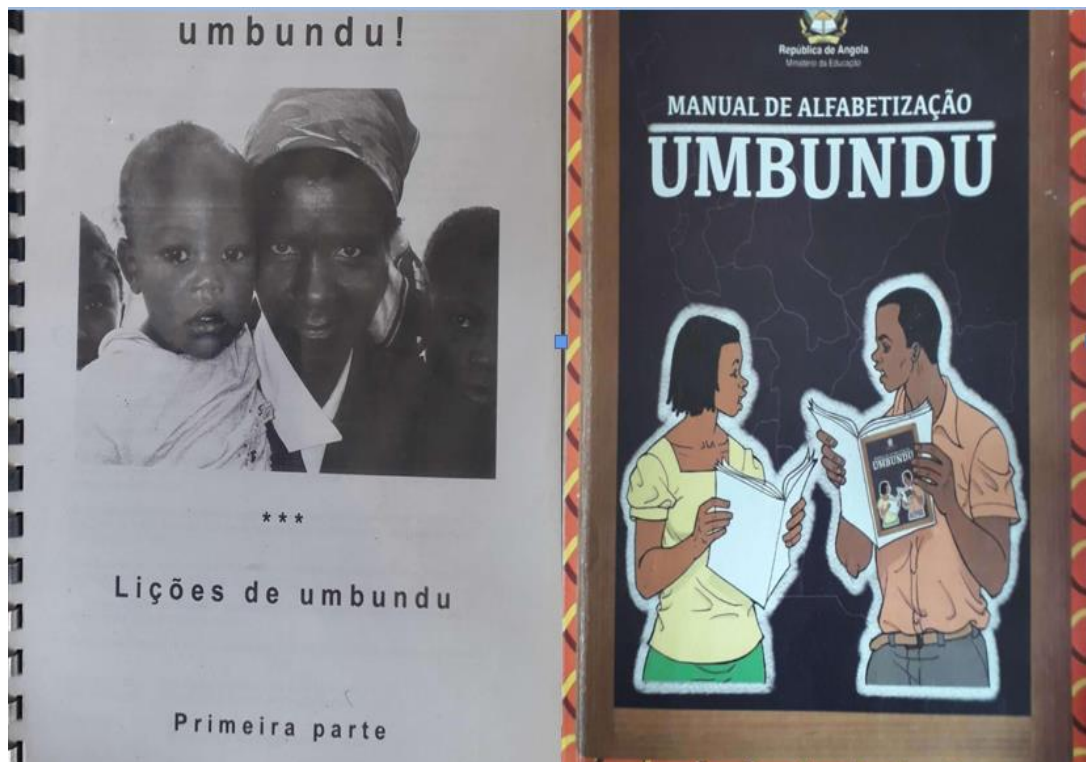


Note. Le Guennek & Valente Umbundu Dictionary (Le Guennek & Valente, 1972) and Etaungo Daniel's Portuguese-Umbundu Dictionary (Daniel, 2010) respectively

Figure 5 shows two dictionaries of Portuguese – Umbundu. The dictionaries have a peculiarity of representing two crucial orthographies. The one on the left was written by Catholic priests, while the one on the right was written by a Protestant pastor. The dictionaries represent a difference in the orthographies for the Umbundu language. There have not been any known conventions, before or after the independence of Angola, on the orthography of this important language. And, this has contributed, to a greater extent, to the existence of different orthographies even in the tests use in State schools at the experimental level.

Apart from dictionaries, other didactic books and leaflets have been proposed for classroom use. Most of the available materials have been designed by missionaries at the beginning and later proposed by the State at an experimental level. In fact, enough classroom resources have been produced under the initiative of the Angolan government, nowadays. Below is a sample of some books which have been made to help learners and teachers in their teaching and learning process.

Figure 6: Didactical material for the classroom



Note. The first Title page is of Protestant missionaries in Caluquembe (n.d.); the second is the State's initiative for public schools.

Figure 6 presents a sample of didactical material proposed for the teaching of Umbundu. The book on the left represents an initiative from non-governmental agency while the book on the right

has been proposed by the State. Again, the two didactic materials represent different orthographical approaches for the Umbundu language. The one on the left was written under the Protestant orthography (the cover, here has been written in Portuguese, meaning “Umbundu lessons – First part”), while the second was written under the orthographical principles adopted by the State (the title States “Alphabetization Manual – Umbundu). The State, as it will be seen in the following chapter, has made a kind of mixture of the PSO and the CSO.

Another attention to pay when designing the orthography of a language is on the percentage of people who are expected to proficiently use the orthography. In this regard, one has to reflect on the number of people who belong to the Umbundu language community which corresponds between 23% and 37% of the total population of Angola. Although studies indicate that the number of speakers decreases drastically as new generations take over, mainly in larger cities of the shoreline, the expectation with regard to the school coverage is high and that there still are many people whose use of Umbundu orthography is vivid on daily basis. Such is the case of priests and pastors who have to read the biblical passages and the doctrine in both Umbundu and Portuguese in their celebrations which occur early in the morning and late in the afternoon in the synagogues. Besides, there are teachers of Umbundu, be they trained or amateurs, whose preoccupation shall be to be proficient in the language and its orthographical rules. These professionals have been working at selected institutions where Umbundu is taught at the experimental level. So, although the percentage of those who are expected to use the Umbundu orthography proficiently cannot be precise, the number of potential users may be high if the number of people who can speak Umbundu, according to the Census described above, is considered.

Finally, it is important to be aware of the attitudes of the various communities constituting the linguistic group towards the orthography to be adopted. As to the users attitudes towards the orthography to be adopted, the information about the two principal orthographical standards (the PSO and the CSO), show that there will be two principal groups: one supporting the PSO and another one spurring for the CSO. Besides, there are those who will follow the State-adopted orthographical approach, let alone the slight variations adopted by various researchers proposing alternative ways of writing the Umbundu language.

b. Execution aspect.

This aspect is related to the display of the letters on the screen, its availability for prompt use in the typewriters or computer keyboards and softwares. For this aspect to be consistent with the

effectivity of language orthography, certain requisites have to be satisfied. The first of them has to do with the inventory of characters, i.e. the letters that will constitute the basic alphabet for the Language. Hosken (2003, p. 4) offers sufficiently clear procedures on the characters inventory:

- List all characters in the alphabet, in what is considered the alphabetical order, if there is one. Many scripts do not include all the characters in the alphabet. Diacritics are rarely included, but they are characters just like base characters. It is, therefore, important to list all the consonants, vowels, tone marks, etc.
- List all case relationships. If an orthography has the concept of case (upper and lower case), then the relationships between characters in the various cases should be listed.
- List all the punctuation characters and their functions. Very often such information is borrowed from another orthography. If so, then this should be stated, along with any differences in this orthography. Are all the punctuation characters borrowed from another orthography, or are some not used?
- List all the consonant sequences and vowel sequences. The consonant sequences correspond to consonant clusters, i.e. we are only concerned with consonant sequences within a syllable. It is not necessary to list all syllable final consonants followed by syllable initial consonants.

The procedures above observe four important items: the first one is the selection of necessary for the orthography to work effectively, the second focus on the distinction between majuscules and minuscules, then comes the punctuation marks. Finally, the sequencing of characters should be defined.

Another important requisite is the effects of the letters when collocating with others, i.e. what consequences, in terms of phonetic-semantic value, will there be when two or more characters are combined? The following guiding questions are crucial when dealing with characters' disposition in the orthography conception:

- For each diacritic, what can it attach to? Include other diacritics, if diacritics can stack.
- Are there any required ligatures? How about optional ligatures? This is particularly pertinent to Indic based orthographies where ligatures are used for conjuncts.
- Are there any particular character shapes this orthography uses? [...] Which one does this orthography use? Or the orthography may call for a script 'a' versus a print 'a'.
- Describe any contextual shape variations. For example consonants in Arabic based orthographies have different forms dependent upon the position of the character in a word (Hosken, *loc. cit*)

The behaviour of characters is worth discussing. The simpler (say, shallower) the orthography the easier it becomes for learning and teaching. Let the figure below be taken:

Figure 7: Sounds produced according to characters' and/or diacritics' combinations

Potência colonizadora	SOM									
	[u]	[i]	[o]	[ʃ]	[ʒ]	[ɲ]	[ŋ]	[tʃ]	[dʒ]	[s]
Inglaterra	<u>	<i> <e>	<u> <o>	<sh>	<zh> <z>	<ny>	<ny> <ng>	<ch> <tch>	<j> <g>	<s>
França	<ou>	<e>	<o>	<ch> <tsh>	<g> <j>	<ny> <ɲ> <ñ>	<ny> <ñ>	<tj> <tch> <tsh>	<dj>	<ss>
Portugal	<u>	<i>	<u>	<ch> <sh> <x>	<zh> <xj> <jh> <j> <g>	<ny> <ng> <nh> <ñ>	<ng'> <ñ> <ḡ> <ñg>	<c> <ch> <tj> <tx> <ty> <tch>	<dj>	<c> <s> <ss> <x>

Note. The figure above, in Jimbi and Sicala (2020, p. 112), shows different orthographical decisions made by each colonial power (i.e. England, France and Portugal, respectively) on African languages' orthographies. As for Portugal, for example, one finds 4 diacritic for a single nasal sound – [ŋ].

Figure 7 is a table showing the way each colonizing power influenced the way African languages were written. As can be seen, 10 phonemic symbols have been presented and their respective alphabetic representations for the local languages of Africa. It is clear from the table that the 3 colonial powers have influenced the writing of the phonemes according to how they write their own languages, be they French, English or Portuguese.

In rigor, the representation of the nasalized velar sound [ŋ], can most probably only find one approximate character: the hard velar <g> which, if nasalized with the diacritic “tilde” – <ḡ> – is subject to represent a nasalized velar sound. Any other attempt to represent it may result in a complexity to conform to. And, to avoid unnecessary complexity, orthographers can decide to revisit other orthographies servicing the languages of the world to better fill in the gap (Hosken, 2003, p. 4).

c. Word splits.

The orthography should indicate where words shall be split. For the Umbundu language, which is based on the Portuguese alphabet, the places of the word part where a string needs to be split by a hyphen shall be made clear for both the learners and the teachers. How syllables are separated when a word does not fit the end of a line should be clarified in a section of the orthographic description. It is justifiable, here, to state that the Umbundu language is written in a way that is complex because the language has not been sufficiently studied for decision-makers to decide which string is a word and which one is really only a bound morpheme. However, in general, “line breaking” in Umbundu tends to be made on the basis of Portuguese orthography.

d. Sorting order

Sorting order is technical expression standing for the logical sequencing of linguistic items that has been decided for the orthography (Hosken, 2003, p. 5). For example, given two words, one being a subject and another being a verb, in English, the orthography demands that the sequence of these strings in the text will be as follows: subject + verb (e.g. “I know”) and not verb + subject (e.g. “know I”). Another example of sorting order is ascendance of numerals (e.g. “1; 2; 3; 4...”, “I; II; III; IV...”) and the ascendance of the alphabet and its specification of lower and upper cases (i.e. “Aa; Bb; Cc...”). Additional example is considering that in the alphabetical orthographies strings of words are written from left to right, while in the Arabic orthography writing texts is made from right to left. All this information needs to be described in the new orthography to be proposed. Sorting orders may be multiples in the orthography.

e. Sample text

A sample text should be provided and it should be translated in the main language of the country as well as in the language in which it is described. For example, a sample text in the Umbundu language under the new orthography should be provided, translated into Portuguese as the main language of Angola and, finally, into English which is the language used to describe the proposed revision of the orthography.

f. Important referential resources

Reference material should be made of as many as possible material as possible. In this list of materials, there should be dictionaries, word lists, various textual productions, phonological descriptions of the language and the like. All this information need to be reasonable and passible to be accepted by the community of the Umbundu readers in the new orthography.

g. The type of scripts

It is important to decide on the type of writing system will be used. Many are the nowadays writing systems: the Roman script, the Indic script, Arabic, the Southeast Asia script, the Syllabary one, or ideographic ones. For our study, the roman script has been chosen to write/harmonize the Umbundu orthography.

h. Computing

Other crucial information to consider is related to computing. Computing is one of the nearest stages towards orthography implementation when programming of characters and other symbols are processed. Nowadays' computers have UNICODE – the universal standard for characters encoding – incorporated as a vital resource to create orthographies (and/or orthographical gaps),

mainly those writing systems of which spoken language articulation finds no characters to correspond more accurately. As a matter of fact, there are sounds in Umbundu language, such as the velar [ŋ] and the palatal [ɲ] whose representations in the Latin writing system are very confusing (cf figure 7). [ɔtʃi 'ɲama] for “animal”, is written “*otchin~~h~~ama*” in the CSO and “ocinyama” in the PSO. Besides, one finds “ot~~j~~ñama” in other books (see de OLIVEIRA, 2012, pp. 909 -1198). As there are no exact characters for such sounds, let alone any convention for the Umbundu orthography to respond to those gaps rigorous and commonsensically, as yet, orthographic revisions have to take this situation into account. The Unicode has been made to allow for problems of this kind.

Special keyboard for a language is essential for the language community to be familiar with. Most of the African languages use European languages keyboards, which makes it difficult to write letters that match with the sounds being produced. This situation may explain the existence of different approaches to writing some peculiar segments by different individuals.

The typesettings available for the orthography of a language is another choke point to take into due consideration. For example, The Portuguese computers typesettings do not include keys for characters, such as, <č >; <ġ >; <ǧ >; <ĵ >; <ñ >; <š >; <ě >; <ı > and <ŭ > (Jimbi & Sicala, 2020). These characters are crucial to write certain sounds without risking being corrupted by means of europenized readings aloud. For example, “Kapiñgālā” (in the CSO), “kapingala”(in notarial services, schools and other official institutions in angola) and “kapiñala” (in the PSO) “kaping’ala” (proposed by Pedro *et al.* in a recent harmonization attempt in 2013) and “kaping~~g~~ālā” (proposed by Valente (1964)) are all to be read [kapi 'ɲala]. This happens because the key setting in European PCs have not been prepared for that. Valente (1964) could go through that easily because typewriters’ typesettings, in that time, could make it. All the aspects discussed above are vital for the development of language orthography.

On the whole, this chapter has discussed the most important criteria and principles guiding (socio) linguists (and other related specialists) in their effort to design or update orthographies for the various languages of the world. It started by discussing the conceptual framework of orthography as an important component of language planning in general and *corpus* planning in particular. Moreover, three basic criteria have been introduced and sufficiently discussed: the phonemic, the etymological and the historical. In this context, an effort has been made to clarify the difference between the often confusing concepts of etymological criterion and that of historical criterion of orthography design. Then, important additional principles have been presented and a link has been established to understand how they feedforward and feedback with one another in an interesting

overlapping manner. The last subsection presents important clues towards the design of orthography for a language.

The following chapter will present and describe the methodological principles underlying the design of this research project by making a description of the objectives of the project, the description of the methodological procedures as well as the structure of the instruments used to approach the data which constitutes the *corpus* to be studied.

4. PERCEPTIONS AND USES OF THE WRITTEN UMBUNDU BY RELEVANT AGENTS

The previous chapter discussed the various aspects related to orthography, and an attempt has been made to establish a link between the theoretical information and the practical development of the orthography for the Umbundu language. It has been clear through the data that there is urgency to engender a reflection on the inclusion of Umbundu into the Angolan education system with caution, given its orthographic situation involving Catholic and Protestant practitioners. This chapter, in its turn, comprehends the use and usage, by relevant agents, of Umbundu. Besides, the chapter explores the awareness, feelings, perceptions, opinions, attitudes, beliefs, convergences, divergences and role of the teachers/trainers in relation to the orthographical standard they choose in the teaching of the Umbundu language in the public schools. Information from teachers and church members on the harmonization of the Umbundu orthography will be collected through the use of questionnaires while document analysis will be carried out to help trace how different the orthographies of the Umbundu language are.

4.1. Methodological approach

This study has been carried out under the triangulation approach (mixed methods). This has been chosen given its flexible nature of allowing for the inclusion of

[T]wo or more sets of data collection using the same methodology, such as from qualitative data sources. Alternatively, the study may use two different data collection methods as with qualitative and quantitative. "This can allow the limitations from each method to be transcended by comparing findings from different perspectives...." (Heale & Forbes, 2013, p. 98).

Bitchener (2010) discusses four categories of triangulation: (1) triangulation on the basis of data when the researcher uses various data for the same study, (2) Triangulation on the basis of the investigators when many investigators are at stake, (3) Triangulation at the level of theory, when various theoretical perspectives are put into use for the research, (4) Triangulation at the level methodology when various methods of data collection instrumentations are involved. Data triangulation and methodological triangulation constitute the most explored kinds of triangulation used for this study.

There are various advantages in using this approach as it may result in convergence, complementariness and divergence of results in the interpretation of the same phenomenon, making

it possible to see how far the findings reinforce the validity of the phenomena, complement each other by showing different aspects of the phenomenon and, even alert to whether a totally new course of research needs to be implemented, respectively (Bitchener, 2010, p. 119).

The methods used for this study encompass the use of questionnaires (one for Umbundu users and the other specific for trainers or teachers) and document analysis (the bibles and other documents showing Umbundu-related corpus). In sociolinguistic research, both field methods (i.e. the use of questionnaires) and analytical methods (i.e. the use of content analysis) are important (Bijeikienė & Tamošiūnaitė 2013, p. 13). And, as Johnstone (2000, p. 36) puts it, “sociolinguistic research projects are neither exclusively quantitative nor exclusively qualitative”. The most important is that the collected data is interpreted and described in a way that it can be well perceived and well understood. The section that follows presents the population, the sample and the sampling techniques.

4.2. Population, sample and sampling techniques

In sociolinguistic research, a population is the group of people that a researcher is interested in or wishes to investigate (Holmes & Hazen, 2014, p. 30; Bijeikienė & Tamošiūnaitė 2013, p. 47). The population of this study is made of the potential Umbundu speaking citizens who also constitute the Catholic and Protestant churches' members, living in the circumscriptions of the municipalities of Benguela, Lobito and Catumbela, in the Angolan province of Benguela. This population is made of participants from various communities where Umbundu is widely used in church services, political rallies, health services, musical concerts, being the mother tongue of the majority of users.

The sample here is defined as the carefully selected and significant representative estimation of the entire population or the typified documentation under a study which can guarantee validity (Rasinger, 2008; Bowen, 2009). On the basis of this understanding, the sample of this study has been made of 64 participants (50 church members in general and 14 teachers/trainers) who have answered the questionnaires between the month of March and May in 2021. These instruments were definitely elaborated and submitted to the final respondents in May of the same year.

To collect data, two sampling techniques have been adopted. On the one hand, a stratified random sampling has been used to deal with the church members (mainly the Catholic and Protestant ones) and teachers or trainers (see Schleeff & Meyerhoff 2010, p. 7). On the other hand, a social network sampling, also known as “snowball”, has been used for its importance in a “friend to friend” approach to data gathering (Tagliamonte 2006; Schleeff Meyerhoff 2010: 7).

Having dealt with the description of the population, the sample and the sampling techniques chosen for this research, an explanation of the instruments used to have the necessary data for analysis will be presented below.

4.3. Research tools

Considering the use of triangulation approach, two questionnaires have been designed as research tools: one questionnaire to be administered to church members and users in general, and another one for teachers and trainers (See Appendixes 1 and 2). Another instrument has been document analysis, including the qualitative information resulting from the written versions of the respondents of the questionnaires submitted to the available Umbundu users (including church members and Umbundu language teachers/trainers). These research tools have been developed to allow for the use of various methodological procedures and generate various inter-supportive perspectives to achieve the purpose of understanding the current nature of the Umbundu orthography for which this study has been devised (see Appendixes 3, 4, 5 and 6). The instruments will be described in terms of how they were structured and their procedural applicability.

4.3.1. Questionnaire on the harmonization of the Umbundu orthography: Design and application

The questionnaire for church members was designed with the purpose of obtaining both quantitative and qualitative data. The title of the questionnaire was “Questionnaire on the harmonization of the Umbundu orthography (Questionário sobre a harmonização da ortografia do Umbundu)”. It was written in English with the Portuguese translation, given within parentheses, to facilitate both English and Portuguese readers.). Given the fact that the present research project pursues the understanding of the nature of the Umbundu orthography, the translation of the questionnaire rubrics into this language did not sound advisable for two reasons: (1) there was no guarantee that the respondents would be competent enough in Umbundu and (2) Umbundu is written in various standards which would make it difficult to choose the one standard to consider for the questions. So the introduction of the questionnaire has explained the objectives of the research and guaranteed anonymity and confidentiality for the respondents. Given the complexity of the study itself, the questionnaire was designed to be an interviewer-administered type (Bijeikienė & Tamošiūnaitė 2013, p. 81) which would provide for the reassurance and certainty that all the questions had been explained, understood, and consciously answered altogether. The reason for following this approach is based on the fact that (1)

some respondents might not be very sure about the questions and (2) the last part of the questionnaire constitutes a group of sentences to be translated into Umbundu, in form of test, to probe the respondents' consciousness of the written standard of their religious confession. A Statement of commitment is left clear at the end of the introduction on the first page of the questionnaire, meaning that respondents who accepted the questionnaires give their consent to participate in the study. Finally, the respondent is appreciated for their answering the questionnaire completely.

The questionnaire has been administered to the respective respondents after a piloting stage. It is important to explain that in social science research, the expression piloting stage refers to "small scale version[s], or trial run[s], done in the course of preparation for the major study" (Polit *et al.*, 2001, p. 467). It is a stage at which the proponent researcher tests the instruments to a minimal number of respondents (Baker, 1994). Among others, one strong point of the piloting stage is its provision of "advance warning about where the research tools could fail... or whether proposed methods or instruments are inappropriate or too complicated" (van Teijlingen & Hundley, 2002, p. 1). Taking the above into account, about 13 respondents were selected among the students of the Department of Modern Languages at ISCED (*Instituto Superior de Ciências da Educação*, translated into English, here, as Higher Institute of Education Sciences)/Benguela. Their feedback was of a great contribution to the final format of the questionnaire (for example, some questions had to be eliminated and the sequence of questions had to be changed to facilitate respondents understanding) that came to be submitted to the final respondents of this study. The final respondents were debriefed prior to responding the questionnaire and assisted as necessary to clarify the meaning of the questions. The respondents were met, contacted and informed about the importance of the study for the making of informed decisions on the Umbundu language orthography towards the necessary harmonization of the Catholic and the Protestant standards for the education system in Angola. As there could be problems with the orthographic standard to be used, respondents were oriented to focus on well chosen words, with the caution to avoid influencing the way they write the words on paper (avoiding bias as much as possible).

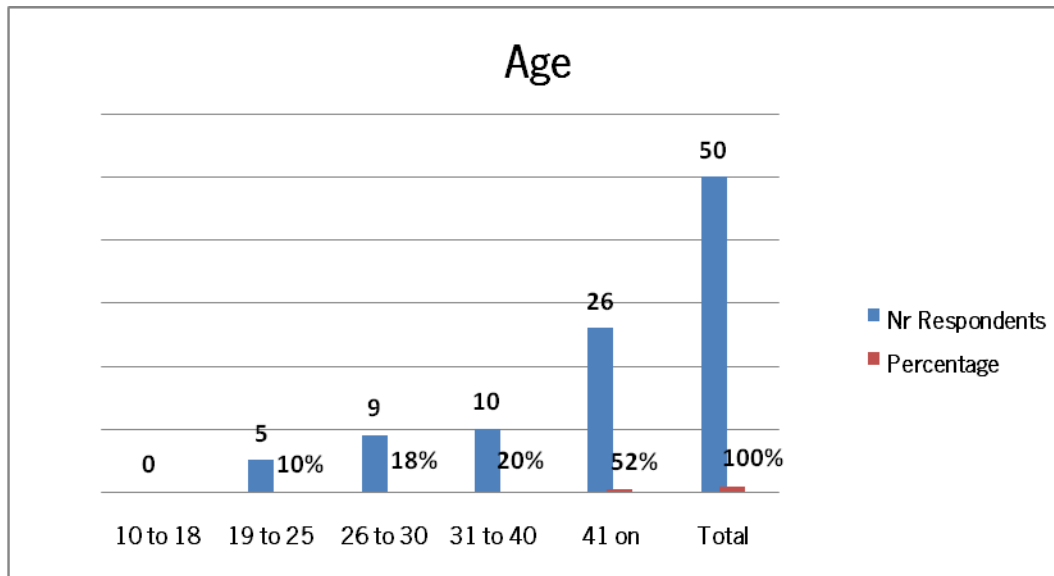
The questionnaire has been designed to have 22 questions split into 4 groups. The first group is about personal information (questions 1 – 10) (see Appendix 1). They are all multiple choice questions intended to capture the respondent's profile of which perceptions about their consciousness of Umbundu, their propensity to learn and use its orthography can be inferred. In this

group of questions, it can be seen whether Umbundu consciousness is more observed within the youngsters or whether it is more practical within senior people.

4.3.1.1. Group one: Personal background

The first question to be considered here is number 2. It has been written in the following way: “Indicate your age according to the categories below. (*Indique a sua idade nas categorias abaixo.*) ”

Figure 8: Age range among respondents



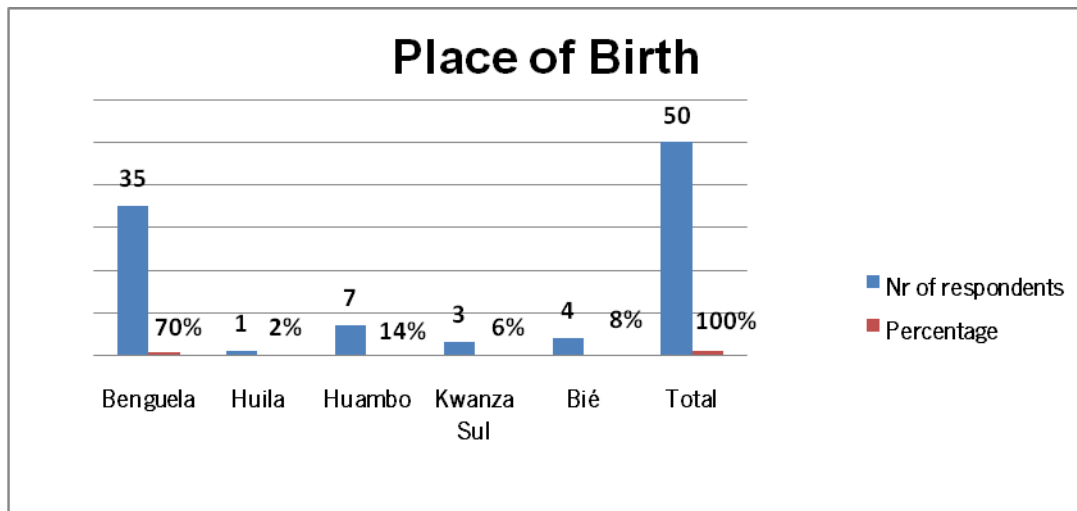
Note. The age range considered was between 10 and more than 40 years old.

Figure 8 shows the age range among respondents. In the graph, the blue bars represent the absolute numbers while the red ones (some are invisible for the percentages are low) represent the relative percentual numbers. The age range has been considered from 10 to over 40 years old. The figures show that of the 50 respondents, the relative majority of 52% are over 40 years old. And none of them is 18 years old or under. The majority of the respondents represent hope with regard to the background support necessary for younger generations to maintain the Umbundu language.

Moreover, it can be known whether the respondents are habitual inhabitants of the places where the research has been conducted or whether they are itinerants. The reason for this group of questions lies in the fact that the language the respondent can speak may be taken as a definite indication of the competence the respondent is expected to have or not to have in Umbundu.

In question 4, respondents were asked to indicate the places where they were born. A graph was drawn to show the statistics as presented below.

Figure 9: Respondents' place of birth

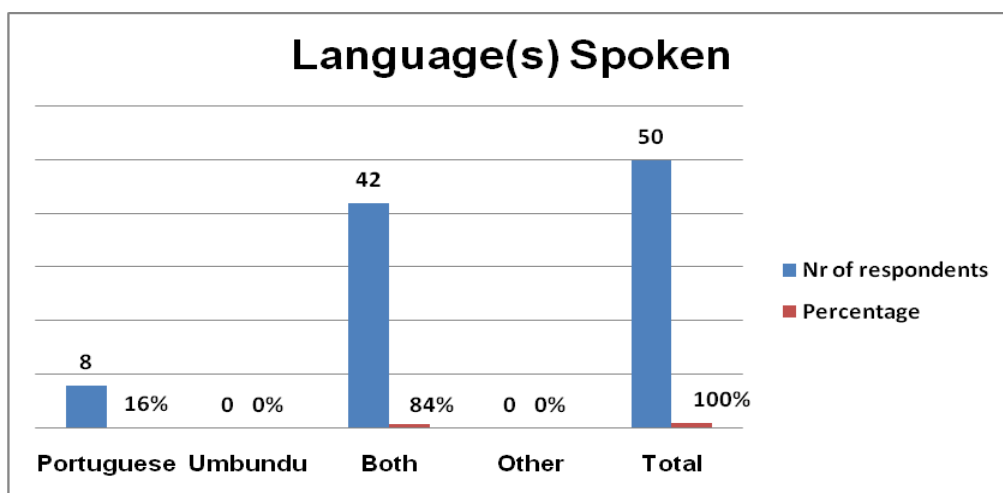


Note. The place of birth is presented considering the provinces where respondents were born.

Figure 9 shows the place of birth of the respondents. The graph comports blue bars for the number of respondents and red bars for the corresponding percentages. It stands clear that a good majority of the respondents (70%) were born in Benguela, 14% in Huambo, 8% in Bié, 6% in Kwanza Sul and 2% in Huila.

Moreover, respondents have been asked to “Indicate the language(s) you can speak. (Indique a (s) língua (s) que falas)” in question 5.

Figure 10: The linguistic repertoire of the respondents



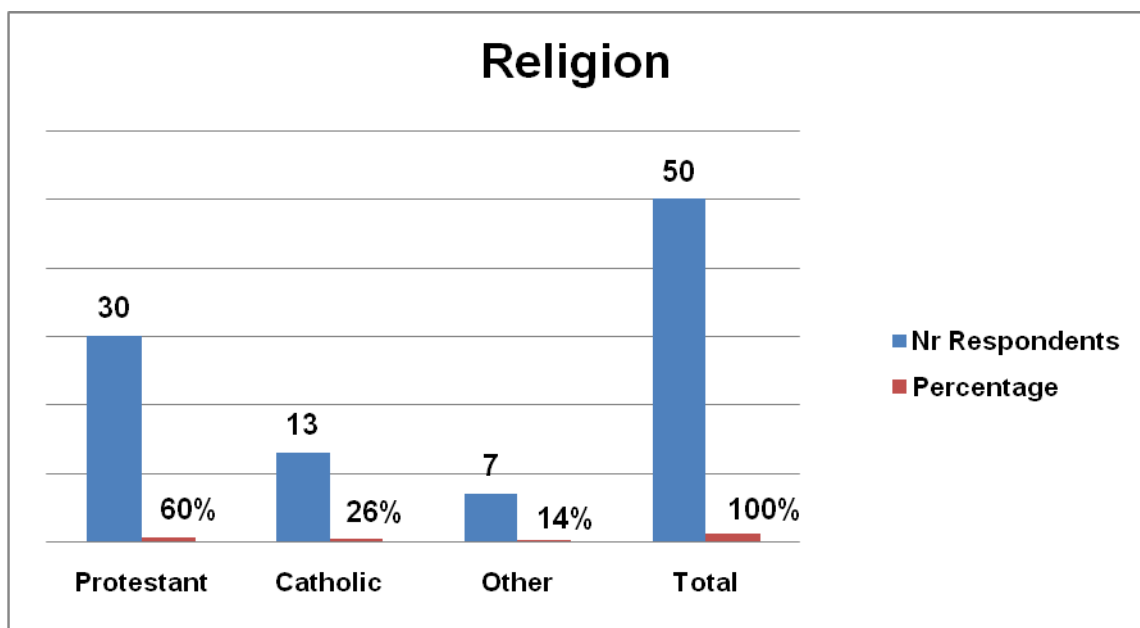
Note. Respondents' linguistic repertoire is important to know for its helping predict the possible language interference.

Figure 10 shows the linguistic repertoire of the respondents. As in preceding graphs, this one has blue bars and red bars. The blue bars stand for the number of respondents and the red bars stand for the percentages. It can be seen that the vast majority (84%) can speak both Umbundu and Portuguese while the remaining 16% can only speak Portuguese.

This group of questions also seeks to reveal the fact that the religion people belong to tells much about the orthographic standard they are expected to use in habitual basis.

The sixth question was about religion: "Write the religious denomination of your own or of your family. (*Qual a tua religião ou a da tua família?*)".

Figure 11: Respondents' religion



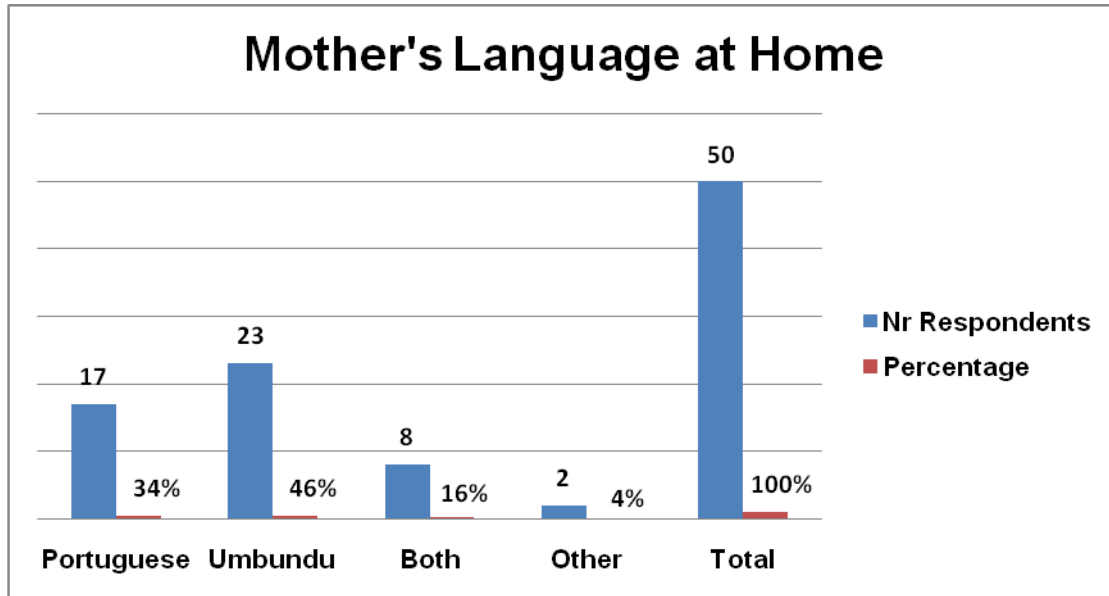
Note. Religion has been the most important preserver of the Umbundu language

Figure 11 shows the respondents' religion. The graph mirrors the statistics about the respondents' religion. The blue bars show the number of respondents and the percentage is shown by the red bars. It can be seen that 60% of the respondents are Protestant, while 26% belong to the Catholic Church. 14% of the respondents were put in the category of "Other" religious denominations.

Finally, this group of questions is important for its capturing the respondents' family's linguistic environment which is crucial to understand the possibility of having or not having instilled the Umbundu language in the respondents.

Question 7 is about the language that the respondent's mother uses at home: "Write the name(s) of language(s) your mother speaks more at home (*Mencione os as línguas que a tua mãe fala mais em casa*)".

Figure 12: Language used by mothers at home

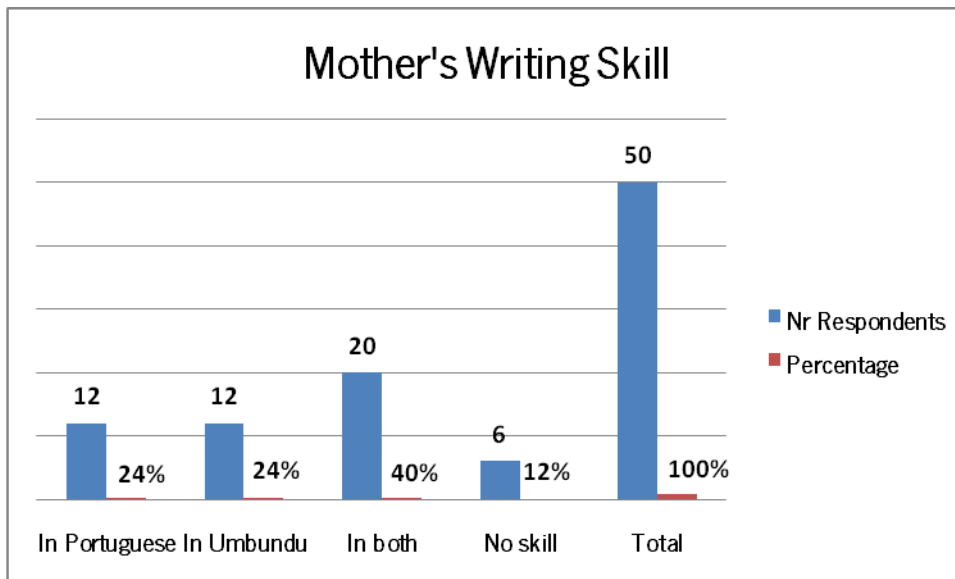


Note. Mothers have a significant linguistic influence on children. Knowing the language they habitually use may help predict how much awareness of the language children may have.

Figure 12 displays the home language used by the respondent's mother. The blue bars show the number of respondents while the red ones are related to the respective percentages. It is clear that 46% of the respondents indicate that the language spoken at home by their mothers is Umbundu while Portuguese is spoken by mothers of respondents at the percentage of 34. 16% of the respondents have indicated that their mothers speak both Portuguese and Umbundu at home. Besides, 2% of the respondents have indicated that their mothers do neither use Portuguese nor Umbundu at home.

In question number 8, respondents were asked to indicate the writing skills of their mothers in the languages they use at home. The graph below has been drawn to represent it.

Figure 13: Mother's writing skill

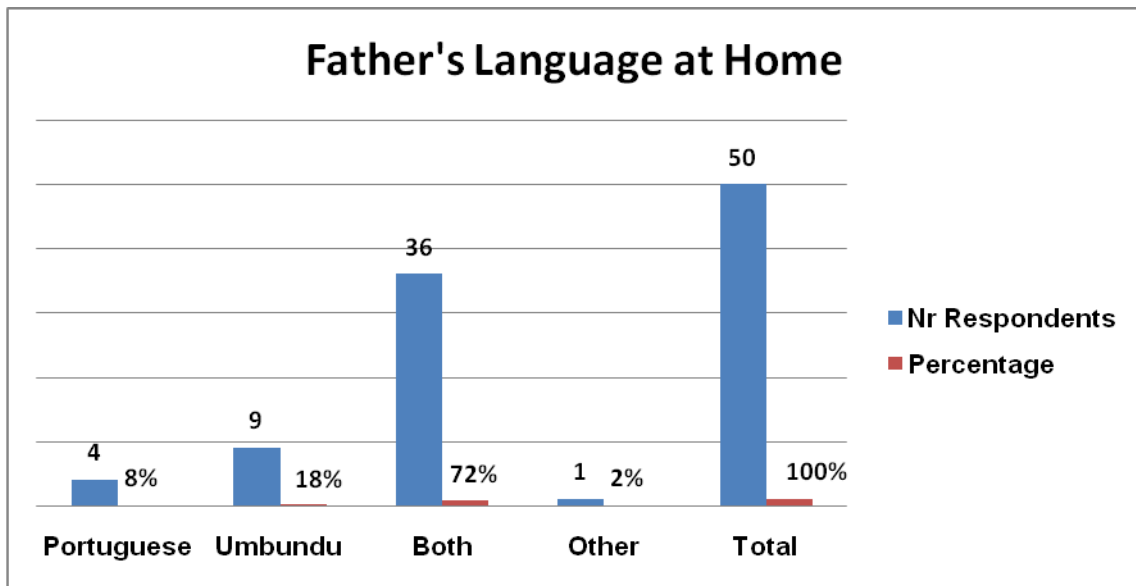


Note: Indication of the existence or absence of writing skills with the respondents' mothers.

Figure 13 shows the mothers' writing skills. As in the preceding graphs, this one is made of the blue bars for the absolute numbers and the red bars for the relative (percentual) numbers. As it can be seen, 24% of the respondents indicate that their mothers can write in Portuguese. A similar percentage can be attributed to those whose mothers can write in Umbundu. On the other hand, 40% of the respondents indicate that their mothers can write in both languages. Only 12% of the respondents have indicated that their mothers have no writing skills in any of the languages.

Question 9 has been designed like this: "Write the name(s) of language(s) your father can speak. (Escreva o(s) nome(s) da(s) língua(s) que o teu pai fala.)".

Figure 14: Language used by father at home

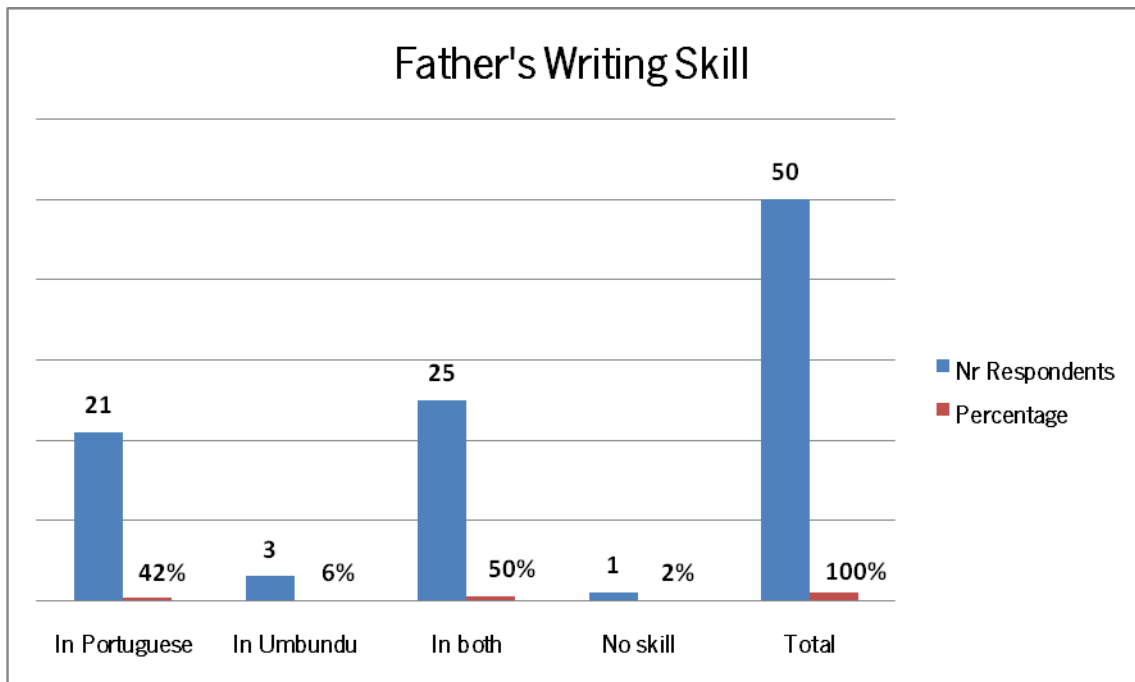


Note. The language used by the father at home, although it may not be as influential as the mother, may determine the pride and the importance given to it by children.

Figure 14 shows the language used by the fathers at home. The blue bars represent the number of respondents while the red bar shows their respective percentages. It is shown that fathers who speak both Umbundu and Portuguese at home correspond to 72% of the respondents. 18% of the participants indicate that their fathers use only Umbundu at home. Only 8% of the respondents have fathers using Portuguese as the only language of communication at home.

Question 10 is as follows: “Indicate whether your father can write in the language(s) that follow. (Indique se o teu pai aprendeu a escrever na (s) língua(s) que se seguem.)”.

Figure 15: Father's writing skills

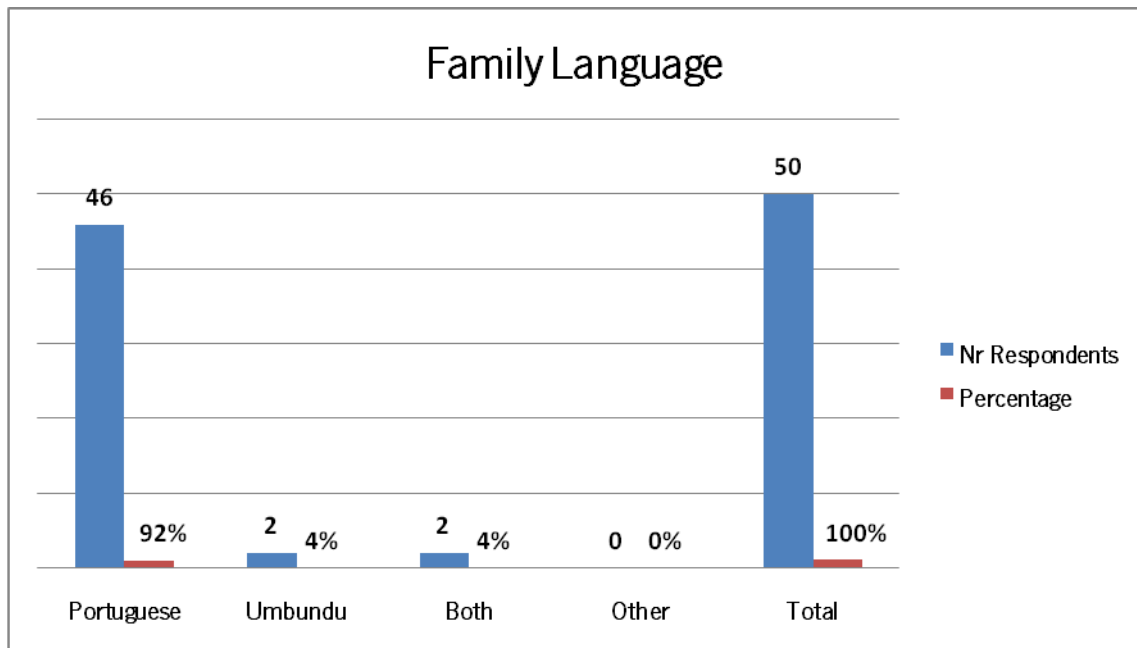


Note. The capability to write may help predict the father's ability to read and write documents, such as letters, reports, catechism and pamphlets. These documents are important inputs for the teaching and learning of language.

Figure 15 shows the fathers' writing skills as indicated by the respondents. The graph shows blue bars for the numbers of respondents and red bars for the percentages. As the graph displays, 42% of the respondents indicate that their fathers can only write in Portuguese while only 6% of the respondents have fathers who can only write in Umbundu. On the other hand, 50% of the respondents (25) indicate that their fathers can write in both languages. Lastly, there is 1 respondent (2%) whose father has no writing skill at any of the languages.

Question 10 asks: "If you have a family, what language do you speak more at home? (*Em caso de viveres em família, que língua falas em casa?*) "

Figure 16: The family language



Note. The language the most used in the family environment may indicate the strength and importance of this language for the members of the family. The less the language is used the more it gets lost in the community that should use it.

Figure 16 shows the data about the language used in the respondents' family in case they have one. The blue bars stand for the number of respondents and the red one represents the respective percentages. It can be seen that a vast majority of the respondents (92%) have Portuguese as a language of their family. On the other hand, 4% of the respondents indicate that Umbundu is the language of the family and a similar percentage of the respondents indicate that both Portuguese and Umbundu are used for the family communication.

4.3.1.2. Conclusive remarks on personal background

The analysis of the data on personal background of the respondents of the "Questionnaire on the Harmonization of the Umbundu Orthography" has made it possible to conclude that (1) the majority of the respondents (72%) can be considered adults for being over 30 years old (fig. 8). These respondents represent a population which was born shortly before and shortly after the Angolan Independence in 1975 whose parents' language was most probably Umbundu – by this time 85% of the Angolan population was illiterate (see Cacumba, 2020: 84). This information transmits some hope of having people who can serve as the vanguard of the Umbundu language and help the young generation to gain consciousness of the language and its importance as a means of communication

on the communities in the outskirts of the larger littoral cities and inside them, in situations such as healthcare, notarial services, court, schooling and the like. Despising the existence of the Umbundu language in these places is giving permission to isolation and exclusion of an important segment of the Umbundu linguistic community. Another importance of age is that younger generations can take advantage of the adults' native knowledge of the language to keep maintain it.

Although not all of them were born in Benguela province (Fig. 9), they may be said to be coming from neighbouring provinces. In fact, Fig. 10 shows that 84% of the respondents have competence in both Portuguese and Umbundu, which makes them eventual bilinguals. Moreover, 62% of the respondents show that their mothers use Umbundu at home, against 50% whose mothers use Portuguese at home, on the one hand (Fig.10). On the other hand, 72% of the respondents indicate that their fathers are bilinguals because they use both Umbundu and Portuguese at home (Fig. 13). If 18% of the respondents whose fathers can only speak Umbundu at home, the number of fathers who speak Umbundu at home increases to 90%.

The percentage of respondents whose parents interact with them in Umbundu at home is encouraging in that it can show that there is a certain level of pride for the language and that there is guarantee that the preservation of the language is still in their agenda. Now, if the Umbundu language teaching is effectively included into the education system, then its use and maintenance will be reinforced. In general, children take examples of their parents, mainly the mothers. So, if the parents show pride of their African languages with their little ones, the little ones will imitate them.

Another important information from the results is found in Fig.14 about the respondents' fathers' writing skills. At least 56% of the respondents' fathers can write in Umbundu. This is a crucial information in that this research work focus on corpus planning which involves the way the Umbundu language is orthographically represented. That is, if parents can write in Umbundu, the probability of gaining from their help with the home work is high and there is a chance to compare the various ways the language has been orthographed.

So, in the main it can be said that the respondents are bilingual and have enough information about the Umbundu language and its various orthographical manifestations according to each one's church membership and literary background.

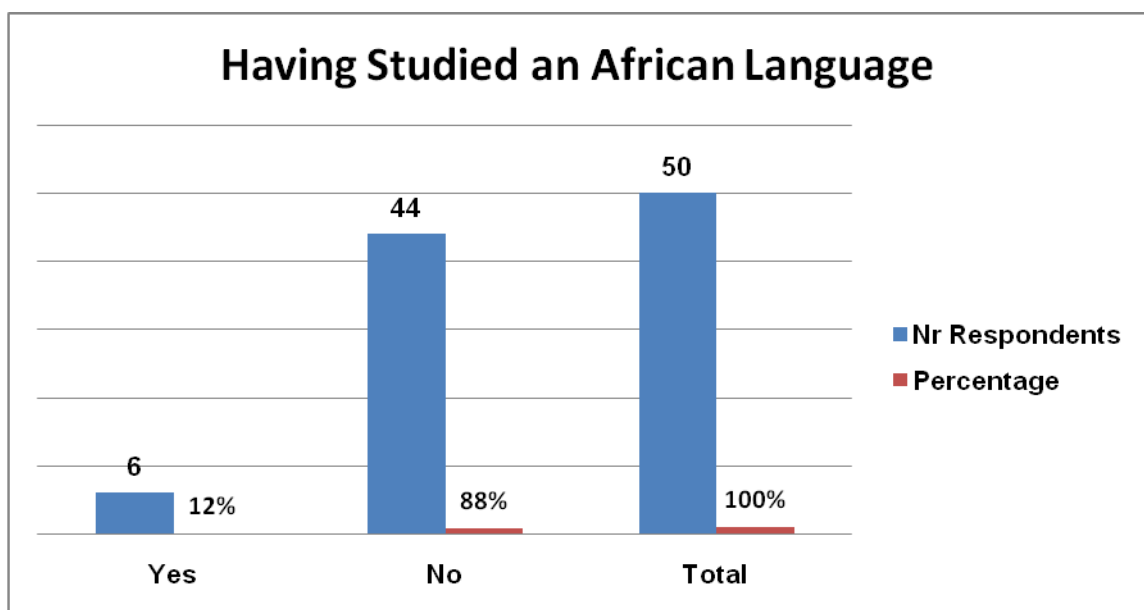
The second group is about language learning which is made of dichotomous and comment questions (questions 11 – 12). The reason for this group of questions lies in its brief capturing information about the situation of the Umbundu language teaching (and other African languages) in the public schools and which public schools exist where Umbundu (and other African languages) is

taught. Another expectation from this group is that the respondents' answer will reveal the standard that they are most inclined to write and read. Also, it gives a very important input with relation to the level of competence the respondents have in the standard they use.

4.3.1.3. Group two: Learning the language.

Question 12 has been written like this: "Have you ever studied an African language of Angola other than Portuguese at public school? (*Já alguma vez estudaste uma língua Africana de Angola que não seja o português na escola pública?*)".

Figure 17: Experience with African languages at school

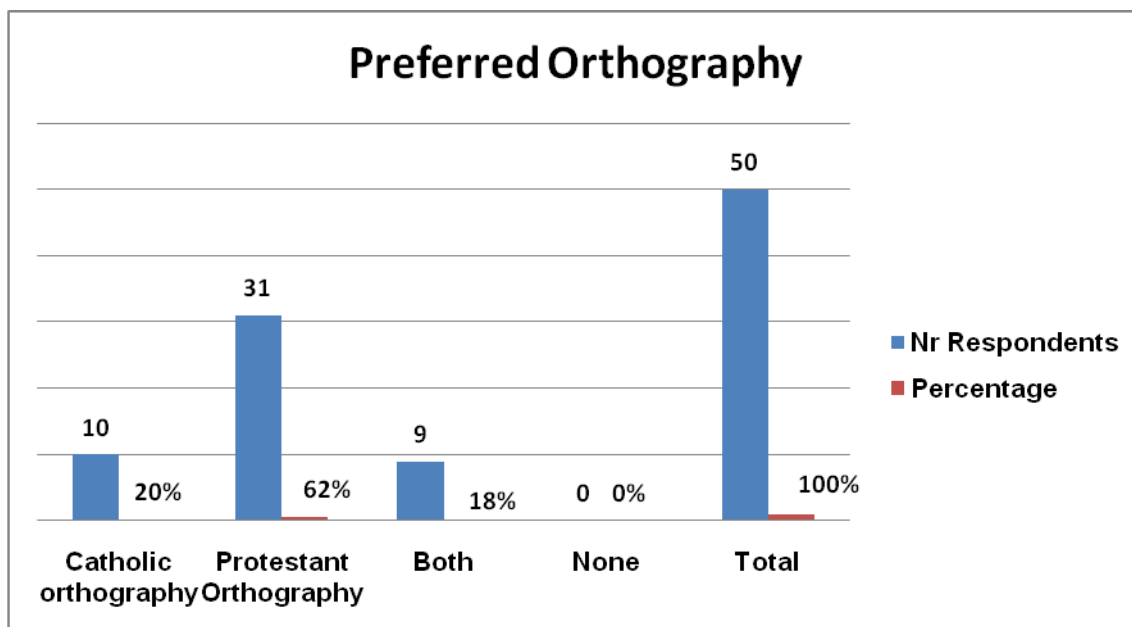


Note. Having experienced the study of African languages in the school context is important information for this study. It indicates the possibilities of existence of teaching programmes, teachers (be they trained, be they untrained and other didactical material for students).

Figure 17 displays data on whether respondents have experienced studying an African language other than Portuguese at school. The graph above shows this information by resorting to a blue bar for the number of respondents and the red bars for the percentages they represent. The table shows that greater majority of the respondents (88%) have not had lessons in African languages at school. However, 12 % of the respondents indicate that they have had lessons in African languages.

Question 12.1. is: “Which Umbundu orthography do you regularly prefer using between the Catholic and the Protestant? (*que ortografia, entre a católica e a Protestante, preferes usar habitualmente?*) ”

Figure 18: Preferred Umbundu orthography



Note. Two important orthographical traditions – the CSO and the PSO – have been influential and the most used in written information of this research.

Figure 18 shows information about the preferred Umbundu orthography of the respondents. The blue bars stand for the number of respondents and the red bars represent the percentages. As it shows, the majority of the respondents (62%) indicate preference for the Protestant orthography. 20% of the respondents prefer the Catholic orthography. Both Catholic and Protestant orthographies are preferred by 18% of the respondents.

4.3.1.4. Conclusive remarks on learning the language

With regard to learning the Umbundu language, very few people have formally studied an African language (12%) against those who have only studied Portuguese (88%), on the one hand (Fig. 16). On the other hand, the majority of the respondents 80% prefer the Protestant orthography (Fig. 18). This information may be an indication that although Umbundu is rare at the public schools, it may be very present at church. Likewise, the data tend to be in favour of the Protestant way of writing Umbundu which may constitute, *per se*, a sign of conflicting position to which standard to validate in the future, if an agreed harmonization, among the most important decision makers and players (i.e. the

Protestant and Catholic teachers and learners, governmental leaders, linguists and other education-related specialists), is not achieved. One important aspect to be considered here is that 60% of the respondents belong to the Protestant church (Fig. 11).

The data above shows that the teaching of the Umbundu language at public schools is not a tradition as it has not been of the experience of the respondents. Besides, it can be advanced that the orthography for the Umbundu language in the education system will still be strongly influenced by two main standards, i.e. teachers tend to teach Umbundu according to the orthography of the religious denomination they belong to as churches have become the current environment in which Umbundu is imprinted in catechesis related literary work and used by the members of the clergy.

As it appears, Umbundu, spoken by 23% of the population as shown in chapter three, has not enjoyed attention of the State, so it could be one of the academic disciplines in the community where it is widely spoken and used as one of the main means of communication and interpersonal interaction. This State of affairs shows that Umbundu has been neglected as a lively tool of interchange within the region that deserves to be taught to guarantee its perpetuation.

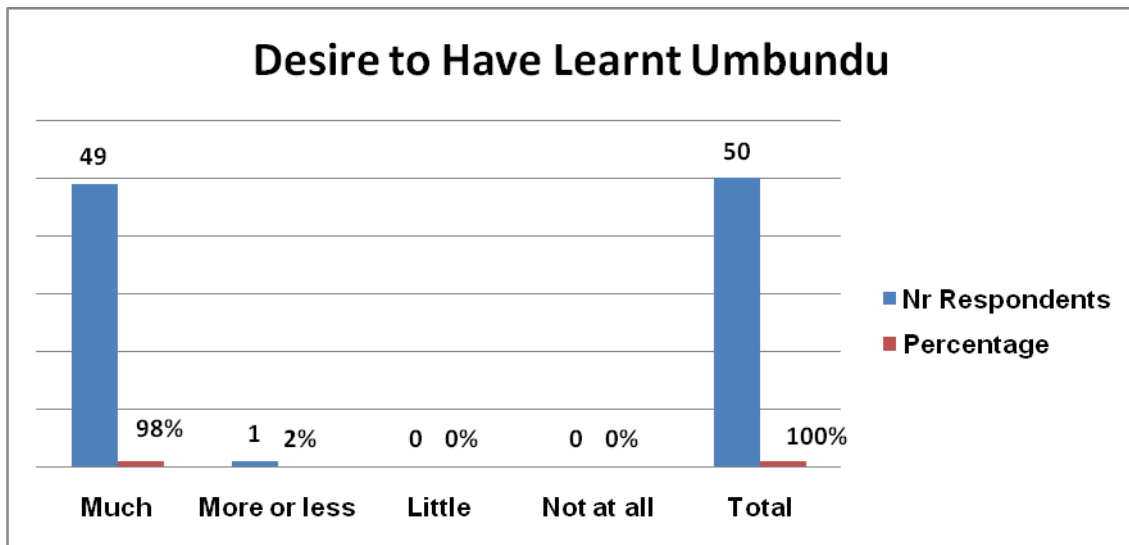
Having dealt with the teaching and maintenance of the Umbundu language both at school and at church, the following section will look at a more practical aspect of Umbundu language by its own users, mainly with regard to the way it is written.

4.3.1.5. Group three: The practice of Umbundu 1

The third group is about the practice of Umbundu (questions 13 – 21). This group is made of multiple-choice questions, Likert Scales, comment and open-ended questions respectively. The justification for these types of questions is that they can reveal the level of coverage, if any, of the Umbundu language studies and teaching in those municipalities of Benguela province (i.e. Benguela, Lobito and Catumbela), as well as give a hint on the general situation of Umbundu studies and the existence of teachers in those municipalities where Umbundu is used as mother tongue and the likelihood of the participants to study Umbundu and to what degree they would like to have learnt that.

Question 14.1. is: “Indicate the degree of concordance with the following Statement: “I would like to have learnt Umbundu at school.” (Indica o grau de concordância com a afirmação... “gostava de ter aprendido o Umbundu na escola...”) ”

Figure 19: Desire to have learnt the Umbundu language



Note. The respondents' willing to have been taught Umbundu or any African language may tell how much pride there is of the language and serve as motivation to insert and generalize it in the country.

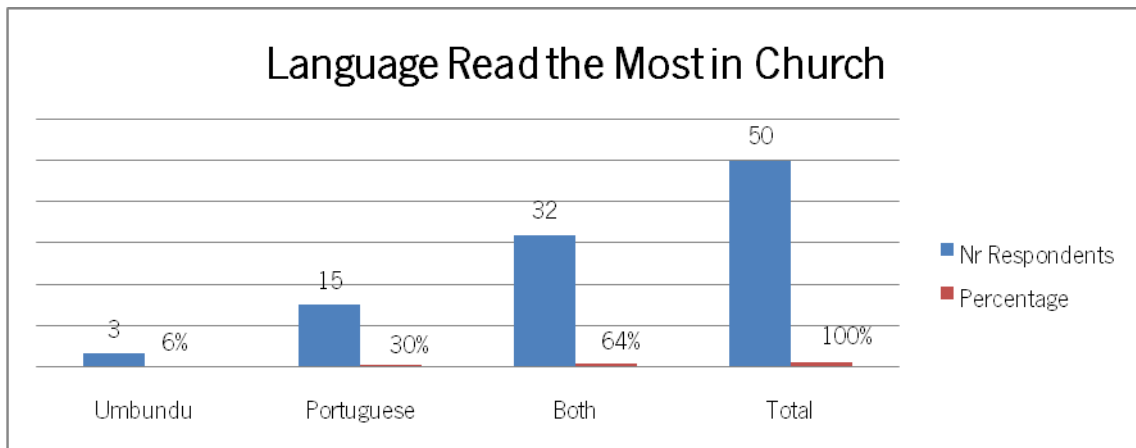
Figure 18 shows a graph of the respondents' degree of desire to have learnt Umbundu at school. The vast majority of the respondents (98%) reveal that they would like to have learnt Umbundu very much. The remaining 2% of the respondents have signaled they would more or less love to have learnt it.

The responses above are very hopeful in that they reveal the respondents' desire to learn the language of their community – Umbundu. When people are motivated to learn a language, they re-energize for it and make provisions to acquire the necessary material and attend lessons with the purpose of gaining enough knowledge to use it for social interaction.

An important point of this group is the revelation of whether the most used language in the church is Portuguese or Umbundu, or whether they are used concomitantly (Fig. 19). The answer to this question also reveals the State of maintenance of the Umbundu language in that specific setting. Moreover, the answers to this group uncover the reading tendencies of the participants in relation to the language in which they prefer to read. In fact, the balance of reading in one or another language may explain whether there is a natural maintenance and support of the Umbundu language. This group is also important for uncovering the reality of the use of a standard orthography of Umbundu in public schools which permits to see how far it differs from those of the Catholic and the Protestant churches.

Question 17 asks to “Indicate the language(s) you prefer reading the most in the church. (*Em que língua (s) mais preferes ler na igreja?*)”

Figure 20: The language used in the respondents' church

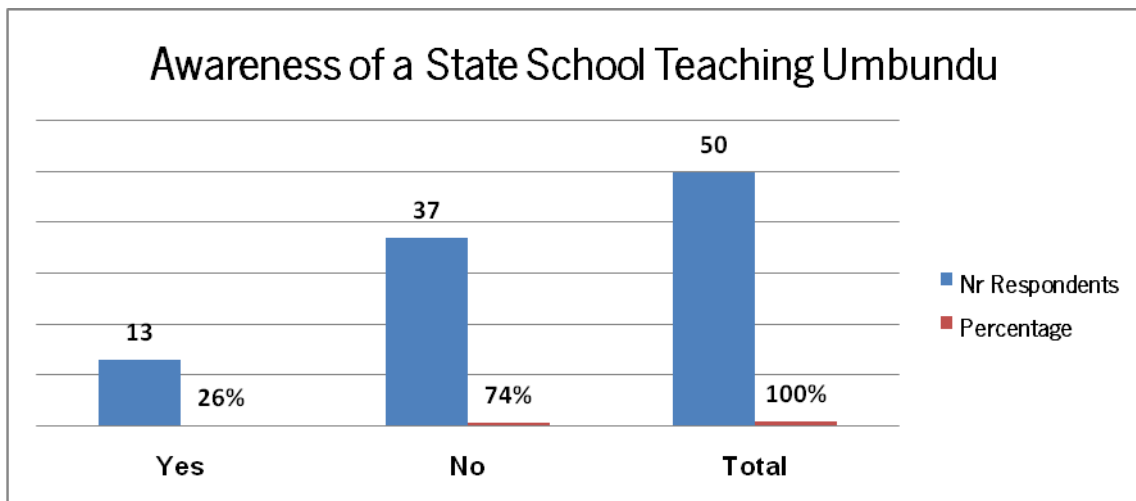


Note. Liturgical work may be done in one language or another. In these congregations, both younger generation and elderly one worship at the same time. They read biblical passages.

Figure 20 reveals the language read the most in church. As in previous graphs, the blue bars show the number of respondents while the red bars show the relevant percentage. Seen from the graph, in the most, both Umbundu and Portuguese are read in the church (64%). Separately, 30% of the respondents indicate that they read in Portuguese, while 6% of the respondents reveal that Umbundu is read in church.

Question 18 is Do you know of a State school where Umbundu is taught? (*Sabes de uma escola pública em que o Umbundo é ensinado?*)

Figure 21: Respondents' awareness of public schools where Umbundu is taught



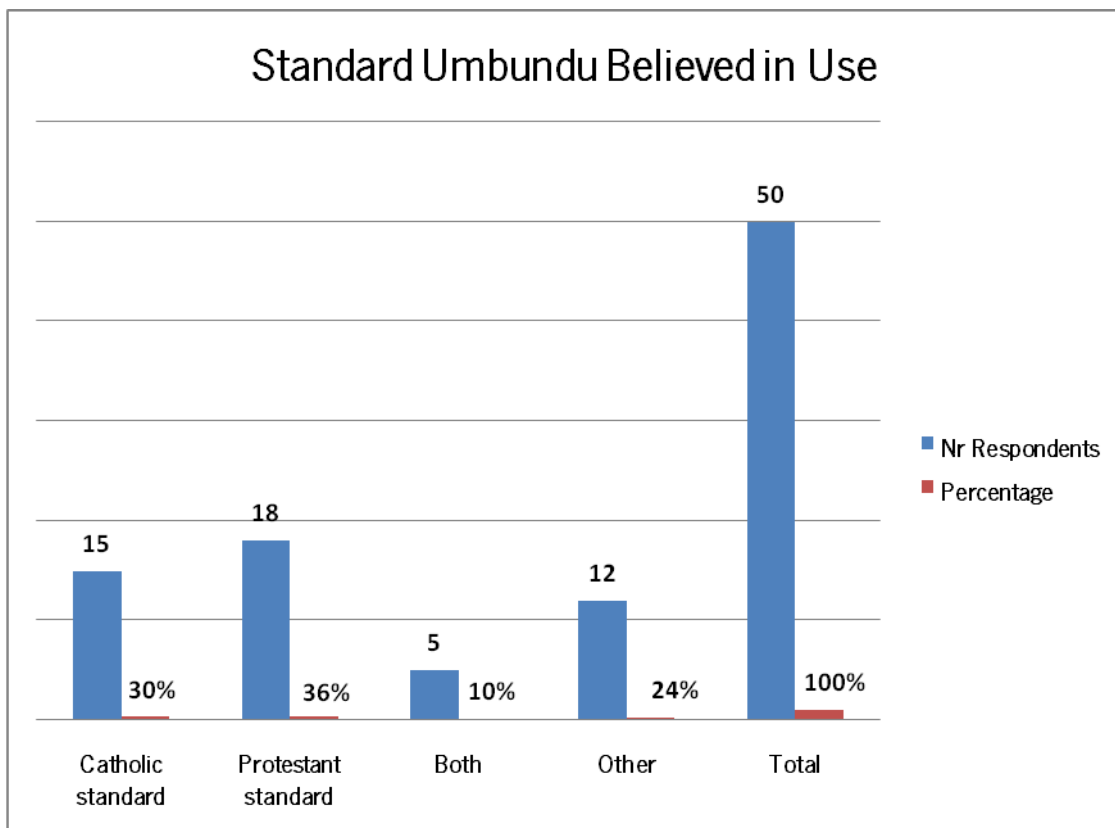
Note. Not every school teaches Umbundu at the experimental level.

Figure 21 shows depicts the respondents' awareness of State schools where Umbundu is currently taught. The blue bars show the number of respondents whereas the red bars show the respective percentage. 74% of the respondents demonstrate awareness of State schools where Umbundu has been currently taught while the other 26% find themselves unaware of any State school teaching Umbundu.

This space here is also important because of the need to capture the respondent's perception and opinion about the standard orthography that could be used. Another important reason of this group of questions is that it gives the respondents an opportunity to comment on the choices they make as there is always a way of learning innovative and creative ideas from comments of the respondent, and shows some tendency for learning Umbundu and if there is a lamentation for not having learnt it when they still were at school.

Question 19 asks "Which Standard Umbundu Orthography is used at State school, do you believe? (Que padrão do Umbundu que acredita estar em usado na escola estatal?) "

Figure 22: Standard Umbundu Orthography that respondents believe that is in use at experimental schools

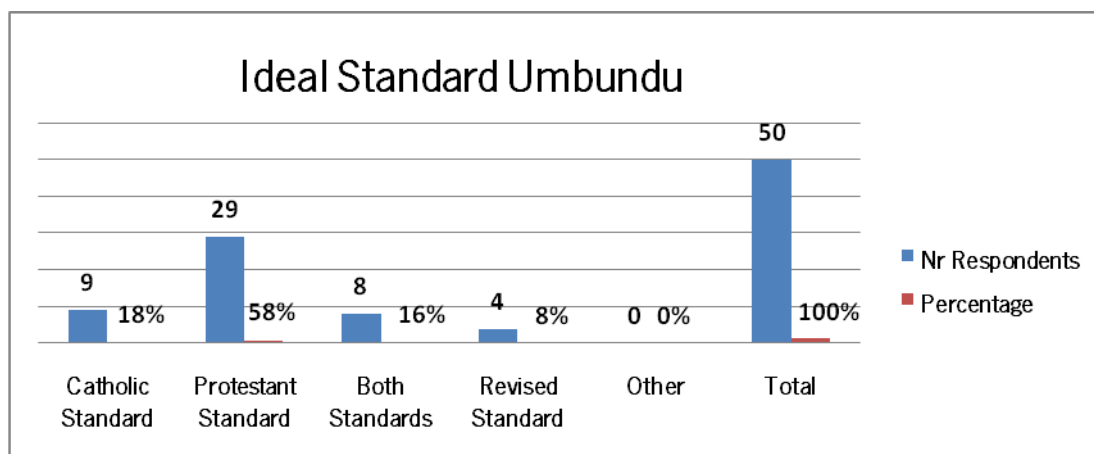


Note. As teachers come from different religious denominations, they tend to teach according to the church they belong to.

Figure 22 shows the standard Umbundu orthography that respondents believe to be in use at schools. The blue bars of the graph stand for the number of respondents and the red bars respond for the respective percentage. It can be seen that 30% of the respondents believe that the Catholic standard is in use while 36% of the respondents believe that it is the Protestant standard that is in use. Besides, 10% of the respondents believe that both the Catholic and the Protestant standards are to be used while other 24% have indicated standards other than the Protestant and the Catholic ones.

A quite similar question has been asked (question 20): “Which Standard Umbundu Orthography should be used at State school, do you believe? (*Que Padrão Ortográfico do Umbundu achas que deveria ser usado na escola?*)

Figure 23: Respondents’ ideal Standard Orthography for Umbundu at school



Note. Although respondents come from specific churches, they may have preferences other than the ones traditionally used in their churches.

Figure 23 shows the respondents’ ideal standard orthography for the Umbundu language. The blue bars of the graph represent the number of respondents while the red bars show the respective percentages. It stands clear that 58 % of the respondents would go for the Protestant orthography of Umbundu while 18% would prefer the Catholic one. Also, 16% would prefer a mix of the two standard orthographies while only 8% of the respondents find it important to choose a revised standard orthography of the Umbundu language.

4.3.1.6. Conclusive remarks on practice of Umbundu 1

The figures above constitute the collection of important data for the understanding of the practice of Umbundu. It can be concluded that the majority of the respondents (64%) have acknowledged that both Umbundu and Portuguese are used for church services (Fig. 19). And, 74% of the respondents do not know any public school where Umbundu is taught, despite the fact that of those who know a

public school where Umbundu is taught (26%) believe that the Protestant orthography is the most used at public schools (Fig.22 & 23). On the other hand, those who choose the Protestant standard orthography for public schools makes 74% of the respondents (Fig. 23). It is important to consider the 8% of the respondents who indicate that the orthography should be revised (Fig. 23).

For the most part, the data suggest that although the general perception of the majority is that Umbundu is not studied in any public school, it is used for church services and many would like to have learnt it at school. As an important language of the communities in that part of Angola, the State should urgently look at the need to teach it as priority, given the fact that as time goes by, many of the people who are experienced in the use of the native languages get old and die in the end. With the demise of the elderly people, the transmission of native knowledge via native languages stops. So, teacher training in Umbundu language teaching shall be found crucial. Also, the opportunity that there is to have Umbundu used in church services should be taken in both hands and incentivated. Most of the people who should be first selected to be trained to teach Umbundu at schools may be the ones who deal with it in churches.

4.3.1.7. Group four: Practice of Umbundu 2 (translations into Umbundu)

The fourth group is made of one question which asks for the respondent to translate sentences, purposefully designed as a test to capture how much the respondent can currently write in Umbundu and which standard orthography the respondent makes use of. This information is also important because the written answer from the respondent can be compared to the Catholic and the Protestant writing systems as well as the various other ways of writing Umbundu that there are.

The question 24 goes that: "Try the most to write the following sentences in Umbundu, so it can be inferred how skillful you are in it: (*Tenta o teu máximo de escrever as seguintes frases na língua Umbundu para que se saiba o quanto podes*)"

Respondents have been asked to translate sentences (see translations attempts in Appendix 5) with the objective of , first, seeing how far they are aware of Umbundu as language of their community, the value that is expected to be given by its users, what they used it for, with what frequency and how they care for it in everyday life; secondly, tracing some of the written patterns mirrored in the translations of the respondents and compare them to both the Catholic and Protestant orthographies, with focus on letter combinations representing specific sounds in Umbundu.

Figure 24: Variations in the written samples done by the respondents

SOUNDS (INTERNATIONAL PHONEMIC ALPHABET)	UMBUNDU WORD'S PHONETIC TRANSCRIPTION/ENGLISH MEANING	SAMPLE WORDS (CATHOLIC TRANSLATIONS)	CATHOLIC STANDARD	NUMBER OF VERSIONS	SAMPLE WORDS (PROTESTANT TRANSLATIONS)	PROTESTANT STANDARD	NUMBER OF VERSIONS
[ŋ]	1. [oŋaŋa] "Baby"	"ngāŋā", "ongāŋā" "nganga", "onganga" and "ŋgaŋga" "oŋgaŋga"	"Oŋgaŋga"	6	"oŋaŋa", "o ŋaŋa", "onana" "nana"	"Oŋaŋa"	4
	2. [onganga] "Witch" or "wizard"	"o nganga" or "onganga"	"Onganga"	2	"Onganga".	"Onganga"	1
[ɲ]	1. [wajɲajɲa] "he/she has sprinkled"	"wanhanha", "wa nhanha", "wanhānhā", "wa nhānhā"	"Wanhanha"	4	"wañiañia", "ya nhanha", "vuanhānhā", "wa nhanhā" and "wa nyanya"	"Wanyanya"	5
	2. [inɲanja] "it uses to defecate"	"yinyanya", "inyanya" and "y nanhanha"	Yinyanya	3	"yiniaŋia", "iniāniā", "iniania", and "ñiañia"	"Iniania"	4
[ʃ]	1. [liʃelena] "it slithers"	"lisielena", "licheleŋa", "lyasielena", "liatchelena" and "isyelena"	"Lisyelena"	5	"liasielena", "licheleŋa", "ichelenā", and "li sielela"	"Lisielena"	4

Note. Respondents have written the same words in different ways. This may signal the lack of writing practice in Umbundu in the respondents' daily life. ²⁷

Figure 24 shows variations in the written samples done by the hands of the respondents themselves. The *corpus* selection in the table above represents solely a small portion of the intricacies that challenge the orthographical arrangements of the Umbundu language. This information shows that people, at times, write in Umbundu with no consciousness of any standard and, some other times, do it with a non-consistent consciousness of a religious standard, be it CSO, PSO or the other. For example, Fig, 23 shows how often the same word varies as it gets written by different people. There is a clear evidence of the lack of training of the respondents with respect to the Umbundu language orthography. Let us take the orthographical representation of the nasal and the nonasal sounds [ŋ] and [g] respectively. The CSO for "baby" varies 6 times and in the PSO changes 4 times. Its contrastive counterpart "witch" varies twice only in the CSO. When it comes to the orthographical representation of the palatal sound [ɲ], the CSO word for "sprinkled" varies 4 times and the PSO varies 5 times. The last conflicting sound is the post-alveolar [ʃ], the CSO word for "slithers" varies 5 times and in its PSO counterpart it varies 4 times.

²⁷ The actual written forms of the figures above can be found in the appendix 4, sample 3 and the appemndix 5, sample 4

4.3.1.8. Conclusive remarks on practice of Umbundu 2 (translations into Umbundu)

The translations were requested to evaluate how far the respondents are aware of the orthographical rules of the Umbundu language. The table that shows the variations in the translations made by both Catholic members and Protestant ones is a demonstration of the current instability of the Umbundu orthography. What the above data tells, immediately, is that there seems to be a high level of illiteracy in the Umbundu language. The orthographical approaches present in the variations seen in the table of figure 24 are an indication that Umbundu has not been used in its writing modality and taught to be one orthographical standard for all its users, although it is currently spoken in the family *millieu*.

The phenomenon above does not only happen to the common user when Umbundu is translated by them. As a matter of fact, even teachers differ in the way they write tests for their students, as evidenced in the tests of the appendices 6, 7, 8 and 9 which are written according to the CSO, as compared to the tests in appendices 10, 12, and 13. This information is testimony of the lack of sufficient professional orientation, professional training and the lack of enough practice in the use of the written modality of the Umbundu language. A strange discrepancy happens when comparing the test of appendix 16 whose first rubric brings “olo ‘prefixos’” (where “olo” is separated from “prefixos”), whereas the corresponding expression in the test of the Appendix 17 is completely written in Umbundu as “olonungilandaka”. As it can be seen, the use of “olo” which appears as a separate word in Appendix 16 appears as a prefix in the Appendix 17. The discrepancies are pervasive all along the tests presented in the appendices 6 through 17 at different linguistic levels.

To illustrate the discrepancies existing in orthographical standards used for the tests, the word “and” in Umbundu tests appears as “kwenda” (Appendices 6, first line of the beginning rubric and 8, second line of the text “Ulonga”, commonly used in the CSO and it appears as “kuenda” in the tests of the appendices 9, 10, 11, 12 and 13, which is a common writing style of the PSO.

In the long run, the evidences are crystal clear that the variations show the instability of the orthography of the Umbundu language as learnt and applied by both Catholic members and Protestant ones, and that there is a clear sign of weak literacy in the Umbundu orthography.

Another questionnaire was designed as there was need to see how teachers or trainers see the situation of the Umbundu language in general and its orthography in particular. The following subsection deals with its design and application.

4.3.2. Teacher/Trainer questionnaire: Design and application

The questionnaire for teacher/trainer was made with the purpose of obtaining important quantitative and qualitative data from respondents who have been trained (or simply have been chosen for being native speakers or because of their knowledge of the language) to teach the Umbundu language (at least at an experimental level). The title of the questionnaire is “Teacher/trainer questionnaire on the Umbundu orthography harmonization (Questionário do professor/formador sobre a harmonização da ortografia do Umbundu)” (see Appendix 2). It was written in English with the Portuguese translation given within parentheses. A translation into Umbundu was found unnecessary for the same reasons explained about the questionnaire above). The objectives of the research were stated in the introduction of the questionnaire and anonymity and confidentiality for the respondents were guaranteed. As the previous questionnaire, this one was made to be a “self-administered” one (Bijeikienê & Tamošĩnaitê 2013, *loc. cit.*).

The questionnaire has been designed to have 17 questions. They are all multiple choice questions, except questions 14, 16 and 17 which are comment ones. The questions were mainly designed to indicate whether those who are teachers or trainers of the Umbundu language teachers have been well informed and formally trained or whether they are only there because they can speak the language which, in turn, gives a picture of whether they can be trusted for the roles they play (as teachers and/or trainers). Secondly, it was found necessary to know about the academic degree because the academic degree may indicate the level of proficiency in doing a certain job, i.e. the higher the academic degree the better position they may be to do the job accurately. It was also necessary to know the specialization because it indicates the area they are more capable to work in, i.e. it is expected to have language teachers for teaching the Umbundu language, although non-language teachers may be found adaptable for the job. The respondents’ religion was asked because, in the main, religion has influenced the emergency of different written standards for the Umbundu language (i.e. the Catholic and the Protestant), and it may help find out whether the respondents write according to the standard of their religion. The fifth question inquires the respondent’s place of birth. As in the questionnaire above, this question was asked to establish the potential influence of the community that the respondents come from and their level of proficiency in Umbundu. The sixth asks for the respondent’s age. The answer to this question is crucial because age may determine whether the respondents have been conservative to the cultural importance of the community’s language, being adults more prone to preserve it than younger respondents whose mother tongue may only be Portuguese (mainly those who were born in the cities. The seventh

question asks for whether the respondent studied Umbundu at the primary school and the eighth question asks about how well the respondent learnt Umbundu. These two complementary questions are relevant in that they do not only bring the use of Umbundu in Angolan State or private schools along the time, but also pursue a direct answer to the level of proficiency the respondents have. The ninth question asks for how the respondent came to know Umbundu to the point of being able to teach it. The reason for the ninth question lies in its force to reveal whether the respondent has had formal training as Umbundu teacher or just adapted to teaching Umbundu to fill in the curricular space. The tenth question asks for where the respondent teaches Umbundu. As some of the respondents may not teach Umbundu, this question establishes a clear comprehension of the specialties as teachers. The eleventh question asks about the length of time the respondent has been teaching. This question is important for didactic reasons, i.e. respondents who have been teaching for longer time may be more capable of dealing with teaching issues in terms of principles and procedures which, in turn, may be in the particular teaching of Umbundu. The twelfth question asks for the written standard of Umbundu the respondent uses. As in the questionnaire above, the awareness of the written standard that the respondent uses may indicate the consistency with which they regulate the written standard they teach the learners. The thirteenth question asks for whose responsibility was it for the design of the Umbundu language teaching programme. This question has been designed because there may be cases in which there is no formal guidance for teachers to teach, making teachers the creators of their own programmes for learners, independent of the Ministry of Education. The fourteenth question was about the type of classroom material in use. This question has been asked as a complementary question to the previous one, because it focuses on the didactical material in use at schools, if any. It also explores the State assistance and the creative and innovative capacity of the respondents. The fifteenth question inquires about the respondent's reasons for defending the teaching of Umbundu. As it is vital to know why people do what they do, this question has been made to understand the reasons behind the respondents wanting Umbundu to be taught at public schools. The sixteenth question is a list of aspects that the respondent is asked to choose if he had to contribute for the harmonization of the Umbundu orthography. These aspects were established because the respondent often has an idea of what should be done to make the Umbundu orthography more functional and practical for the learner in the public schools (i.e. the respondent may want to see one standard to be chosen in detriment of the other, or propose innovative ways to overcome insufficiencies. The seventeenth question was to comment on the choice made in the previous question. The reason for this question is to give freedom for the

respondent to comment beyond the limits of the list and give substance to their choices they have taken. The question 18 asks for any additional comment that the respondent might have found important in dealing with the Umbundu language teaching process. Slightly different from the previous question, this is an open question for the respondent to give additional comment, because more often respondents may find themselves full of other ideas which the questionnaire designer has not addressed and which can be of substantial importance for the research.

The questionnaire has been administered to the respective respondents after a piloting stage. As in the above questionnaire 3 lecturers of the Department of Modern languages of ISCED/Benguela accepted to fill in as respondents of the pilot stage of the questionnaire. Their contributions as experienced people with regard to research instruments made it possible to gain additional ideas and substantial corrections in the format and the linguistic imprecision. This way, they contributed to the current State of the final questionnaire. As expected, the target respondents were debriefed prior to answering the questionnaire and assisted where necessary to clarify the meaning of the questions. The respondents were met, contacted and informed about the importance of the study for the making of decisions on the Umbundu language orthography towards the necessary harmonization of the Catholic and the Protestant standards for the education system in Angola.

Another instrument was designed as there was a felt need to analyse existing documents that could be used as evidence of the existence of more than one way of writing the Umbundu language. The following subsection deals with the document analysis, its design and application as a crucial component of this research work.

This chapter has versed on the description of the methods, population, sample, sampling techniques and instruments used in this research. It has been clear that a mixed method (i.e. triangulation) was chosen to support the way the data was collected. So, the population of this study has been divided between those who responded to the “Questionnaire on the Harmonization of the Umbundu Orthography” and the ones who did it to the “Teacher/Trainer Questionnaire on the Umbundu Orthography Harmonization”. It was also used a documents analysis approach to comprehend the orthographical nature of the texts presented in extracts from both the Catholic and the Protestant bibles. Moreover, the sample, the sampling techniques and the instruments have been presented with enough detail for the reader to follow. All the aforeStated has contributed to the comprehension of the reasons, principles and procedures that guided the construction of the relevant instruments used and the justification of the use that was made of them. Complementarily, the

chapter that follows presents the research data in form of tables, graphs and written samples followed by their respective conclusive remarks.

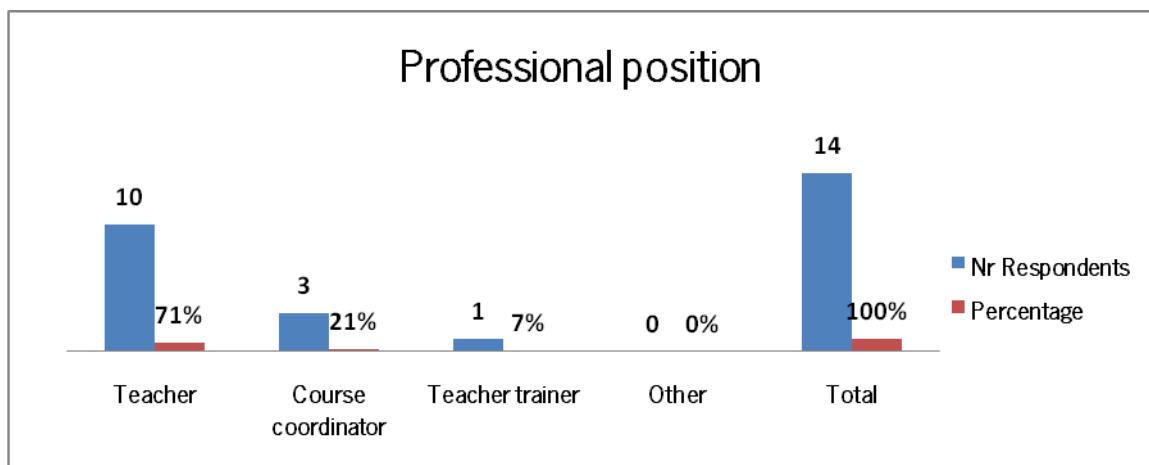
4.3.2.1. Data from the teacher/trainer questionnaire

This section presents the data from the teacher/trainer questionnaire of which essential details are presented above. In effect, the data will be presented as subdivided into three different groups. Data presentation of each group will be followed by its conclusions.

4.3.2.2. Group one: Personal background

Question 1 for the teachers is “What is your current professional position? (Please, circle one option). (*Qual o teu estatuto profissional? Por favor, circule a alínea a que te enquadas*)”

Figure 25: Professional position of the respondents

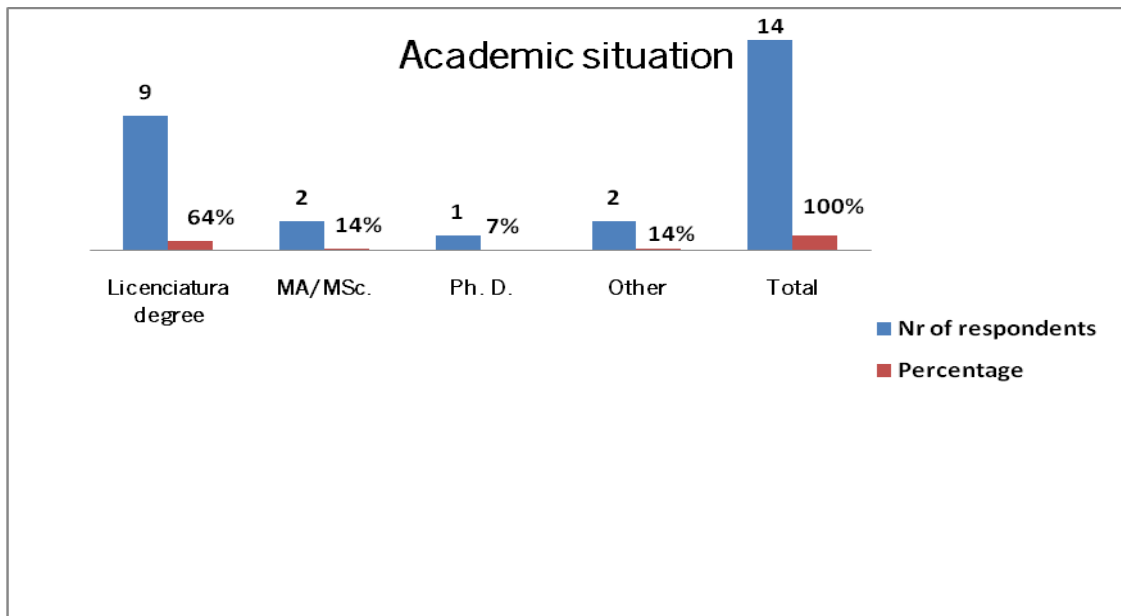


Note. The professional position of the respondents helps predict the influence each element may have and the competences that can be expected from them with regard to Umbundu.

Figure 25 shows the professional position of the participants. The graph displays the professional positions of the respondents of this study. The variables considered in this question include the teachers, the course coordinators, the teacher trainers and other professionals, just in case. As it can be seen, of all the 14 respondents, around 71% are teachers, 21% are course coordinators and 1 respondent (7%) is a teacher trainer.

The question 2 is “What is your current academic situation? (Please circle one option) (*Qual a tua situação académica actual? (Por favour, circule uma opção)*)”.

Figure 26: Academic situation of the respondents

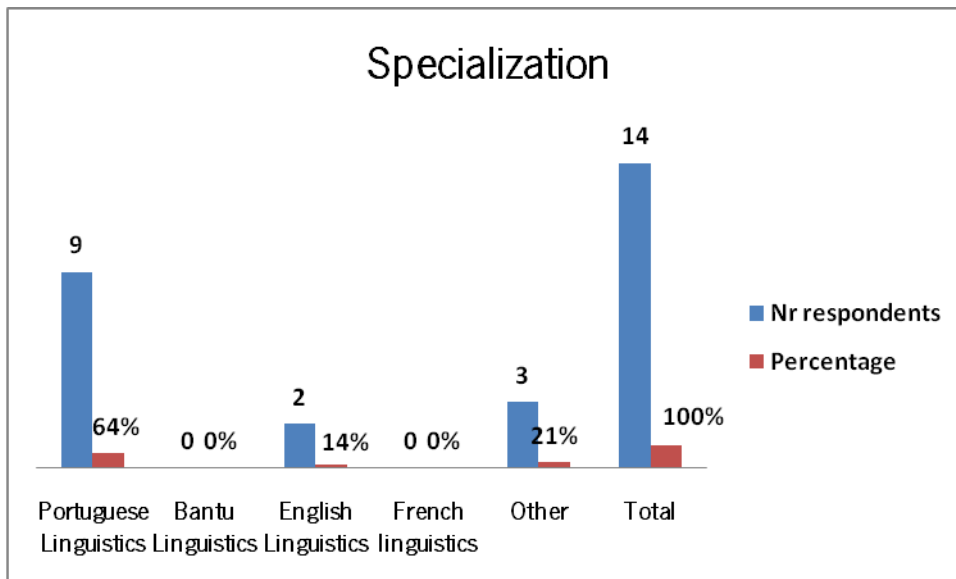


Note. The respondents' academic situation relates to the certifications that they have.

Figure 26 displays the academic situation of the respondents. The graph shows the academic status of the respondents. The blue bars stand for the number of respondents while the red ones correspond to the percentages. It is stands clear that 64% of the respondents have a *Licenciatura degree* (a university degree achieved by students in Angola at the end of a 4-year course). 14% of the respondents have finished their MA/MSc. courses and another 14% have other academic degrees. Only one of the respondents corresponding to 7% has a Ph D degree.

The third question is "What specialization(s) have you done so far? (Please circle one option) (Qual/quais a/as tua/tuas especialização/especializações actualmente (por favor, circula uma opção))"

Figure 27: Respondents' area of specialization

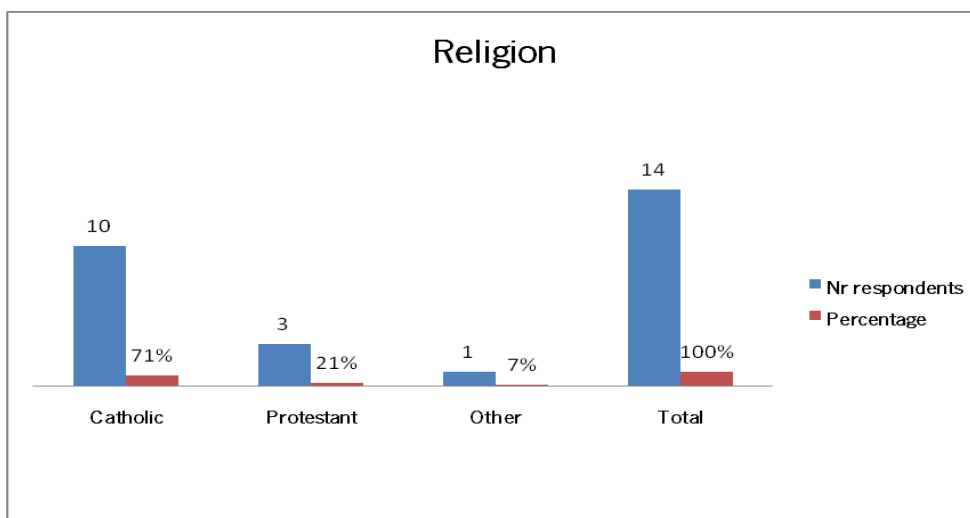


Note. The respondents' area of specialization tell much about whether teachers have been trained to teach Umbundu or solely adapted to fill in spaces of a curricular discipline.

Figure 27 represents the respondents' area of specialization. The graph displays the red bars representing the percentages and the blue bars for the number of respondents. As the it shows, 64% of the respondents have studied Linguistics in Portuguese, 14% have studied Linguistics in English. 21% of the respondents have done other specializations.

The fourth question is "What is your religion? (Please circle one option) (*Qual a tua religião?* (*Por favour, circula uma opção*)).

Figure 28: Respondents' religion

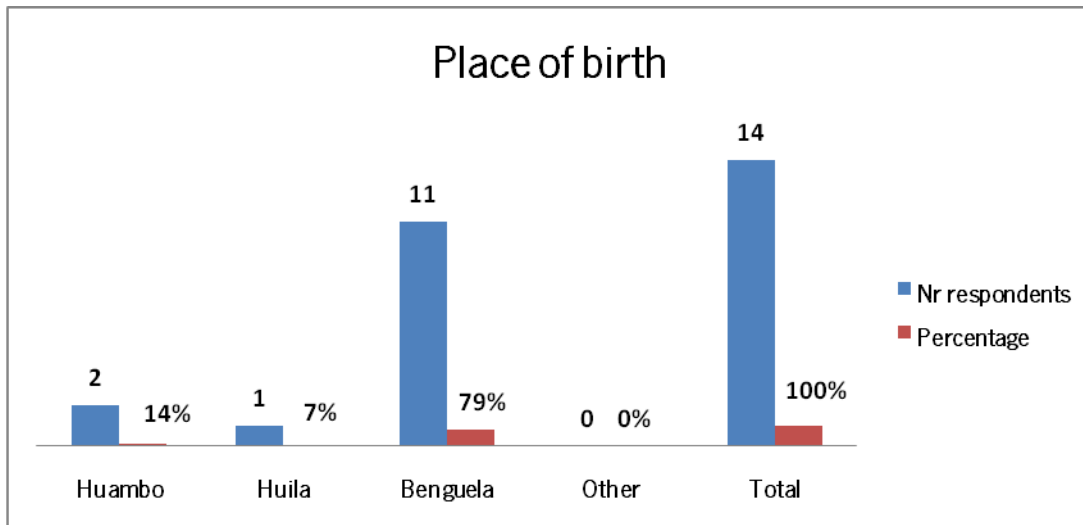


Notes. Respondents' religion can indicate the expectable standard orthography they are prone to use.

Figure 28 displays the religion of the respondents. The number of respondents is shown in blue bars while the respective percentages are presented in red bars. It can be seen that 71% of the respondents belong to the Catholic Church, 21% belong to the Protestant Church and the remaining 7% belongs to other religions.

The fifth question is worded in the following way: Where were you born? (Please, circle one option) (*Onde é que nasceste? (Por favor, circule uma opção).*)

Figure 29: Respondents' place of birth

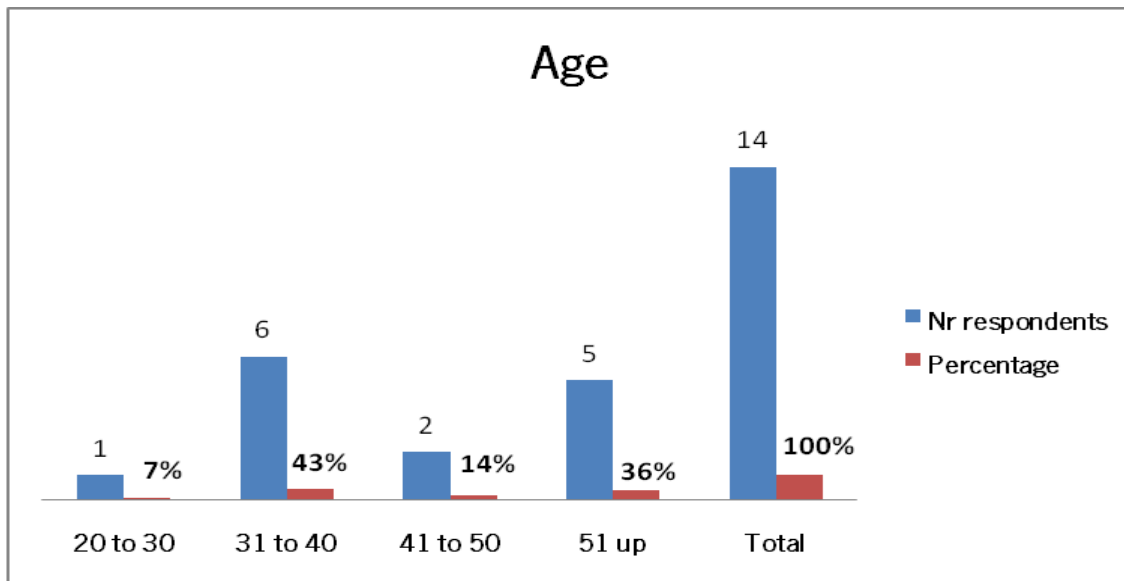


Note. Respondents' place of birth can be linked to the mother tongue or the influence of the language or languages spoken and/or written by the community or region.

Figure 29 evinces the places where respondents were born. The graph above is about the place of birth of the respondents. As in the previous graphs, this one is made of blue bars for the number of respondents and the red bars for the respective percentages. As it can be seen, 79% of the respondents were born in Benguela province, and 14% were born in Huambo province. Only 1 of the respondents, representing 7% of this sample was born in Huila province.

The sixth question is "Choose a category that corresponds to your age. (Please circle one option) (*Escolha a categoria da sua idade. Por favor, circule uma opção*)"

Figure 30: Respondents' age



Note. Respondents' age can indicate whether they have interacted with elderly people making the grassroots behind the conservation of African languages such as Umbundu.

Figure 30 shows the respondents' age. The graph above shows the age of the respondents. The blue bars of the graph show the number of respondents and the red bars stand for the respective percentages. As it can be seen, 43% of the respondents belong to the category of those whose age ranges from 31 to 40. Those who are 51 years old up make 36% of the respondents. Also, 14% of the respondents range from 41 to 50 years old. Only 7% of the respondents range from 20 to 30 years old.

4.3.2.3. Conclusive remarks on personal background

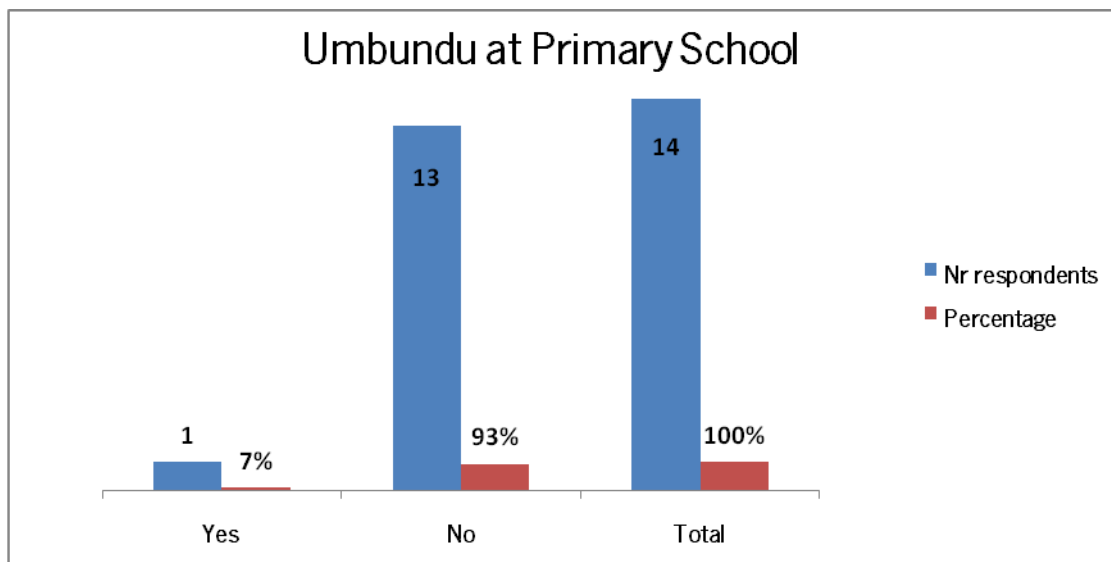
Concerning the personal background of the respondents (teachers/trainers) the majority is made of teachers (71%) with trainers constituting 7% of the participants (Fig. 24). 64% of the participants have finished their *licenciatura* (i.e. a degree awarded after four years in an institution of higher education, corresponding to Bachelor degree), 14% have finished their MA and 7% of the respondents holding a PhD (Fig. 25). Of all the respondents, 64% have specialized in Portuguese Language Teaching (PLT), 14% have specialization in English Language Teaching (ELT) and none has specialized in Bantu Languages Linguistics (BLL) (Fig. 26). Additionally, the majority of teachers/trainers (71%) profess the Catholic faith against 21% of those who profess the Protestant one (Fig. 27). The majority of the respondents are found between the age of 30 and the age of 50.

On balance, it can be safely stated that in terms of teaching skills, there is enough assistance for the classroom given that most of the teachers have sufficient academic level, although the majority have specialized in PLT (Portuguese Language Teaching) and ELT (English Language Teaching). However, there is no specialization in BLL which can guarantee that the knowledge of Umbundu is articulated to its scientific study which, in turn, can contribute to a more academic harmonization of the Umbundu Orthography for the education system on the basis of specialized and more profound debate and literature about the Bantu linguistics. Moreover, the fact that the majority of teachers/trainers are Catholic Church members may mean a future conflict with Protestant Church members in the choice of a standard orthography. Finally, the age of the respondents secures the maintenance of the Umbundu language at least at the level of its orality, an important prerequisite to guarantee that learners are in good hands and that teachers will teach them with native-like language model.

4.3.2.4. Group two: learning the language

The seventh question is “Did you learn Umbundu when you were at the primary school? (Please, circle one option). (*Estudaste Umbundu na escola primária? Por favor, confirme uma opção*)”.

Figure 31: Experience with Umbundu at primary school



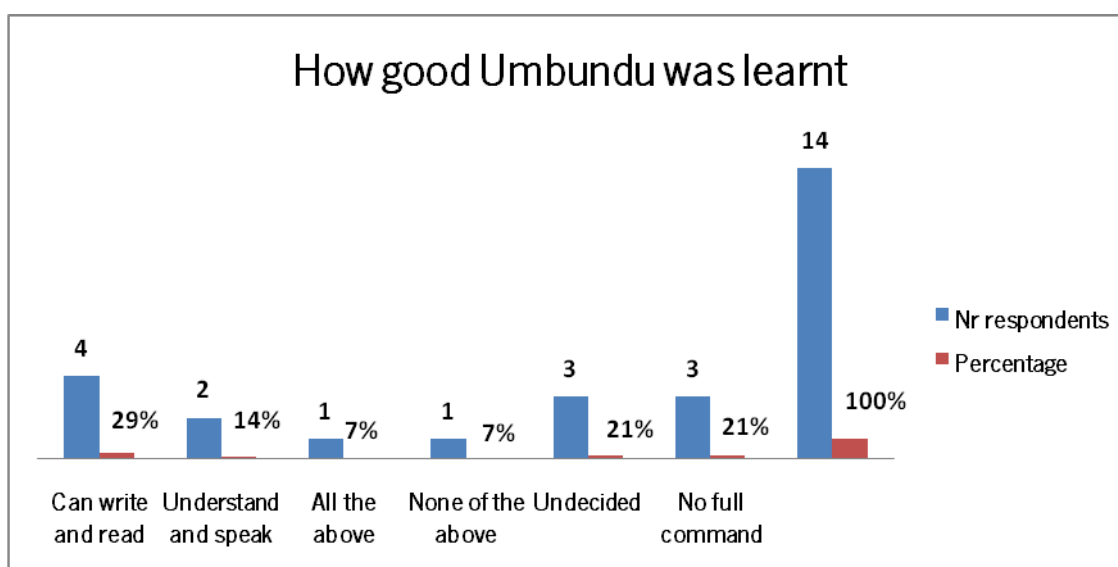
Note. Umbundu as learnt at the primary school can be a sign of an idea of the didactic material and a perception of a standard orthography of the Umbundu language.

Figure 31 presents the respondents’ experience with the Umbundu language at the primary school. The blue bars represent the number of respondents while the red bars represent the

respective percentage. As it can be seen, 93% of the respondents did not learn Umbundu in the primary school. Only 7% (one respondent) indicate that they had Umbundu classes at the primary school.

Question 8 asked “If yes, how well did you learn it? (Please, tick one option) (Se sim, quão bem o aprendeste? Por favor, confirme uma opção) ”

Figure 32: Proficiency and skills in Umbundu

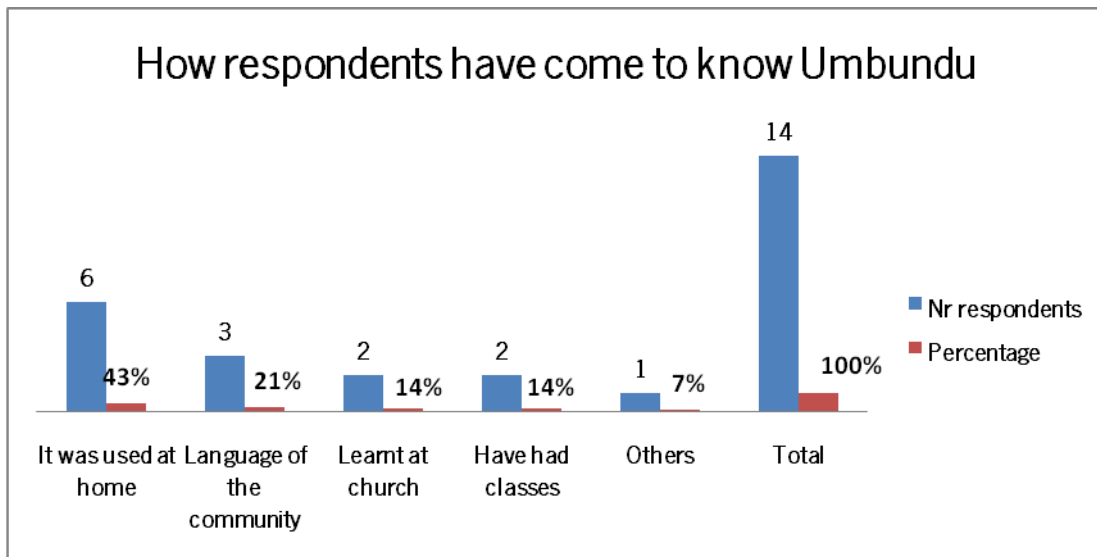


Note. How good Umbundu was learnt in different skills is very important, mainly when it comes to profile candidates to be teachers of this language.

Figure 32 is about how good Umbundu was learnt. The blue bars stand for the number of respondents whereas the red bars represent the percentages. The variables taken into account here include the ability to read Umbundu, understanding and speaking, the domain of the two previous abilities, lack of domain over the two abilities, the ones that cannot be decided and the variable related to the ones who, although they have some awareness of the Umbundu language, they cannot be said to have a full command of it. It can be seen that 29% of the respondents can write and read Umbundu, 14% can understand (orally) and speak it. 21% of the respondents State that they have no full command of Umbundu. Other 21% of the respondents showed indecision about the question.

Question 9 was “If not, how did you come to know it to the point of teaching the future teachers? (Please circle one). (Se não, como é que chegou a conhecê-lo até poder ensiná-lo? Por favor, circule uma opção”).

Figure 33: How the respondents came to be able to know Umbundu to the point of teaching it.



Note. How respondents have come to know Umbundu for teaching it gives an idea of the quality of Umbundu teaching that they can offer. For example, a person whose skills were learnt at home or in the community may be more ready to teach it than a person who learnt it as a school discipline for a few years.

Graph 33 evinces how the respondents came to be able to use Umbundu to the point of teaching it. The graph in the figure above presents the data about how respondents came to know Umbundu to the point that they can teach it to learners. The blue bars show the number of respondents while the red ones show the percentages. As it can be seen 43% of the respondents have come to know Umbundu because it was used at home, 21% learnt it as members of the community, 14% learnt it at church and 14% others learnt it in the classroom. Only 7% learnt it in a non-specified way.

4.3.2.5. Conclusive remarks on learning the language

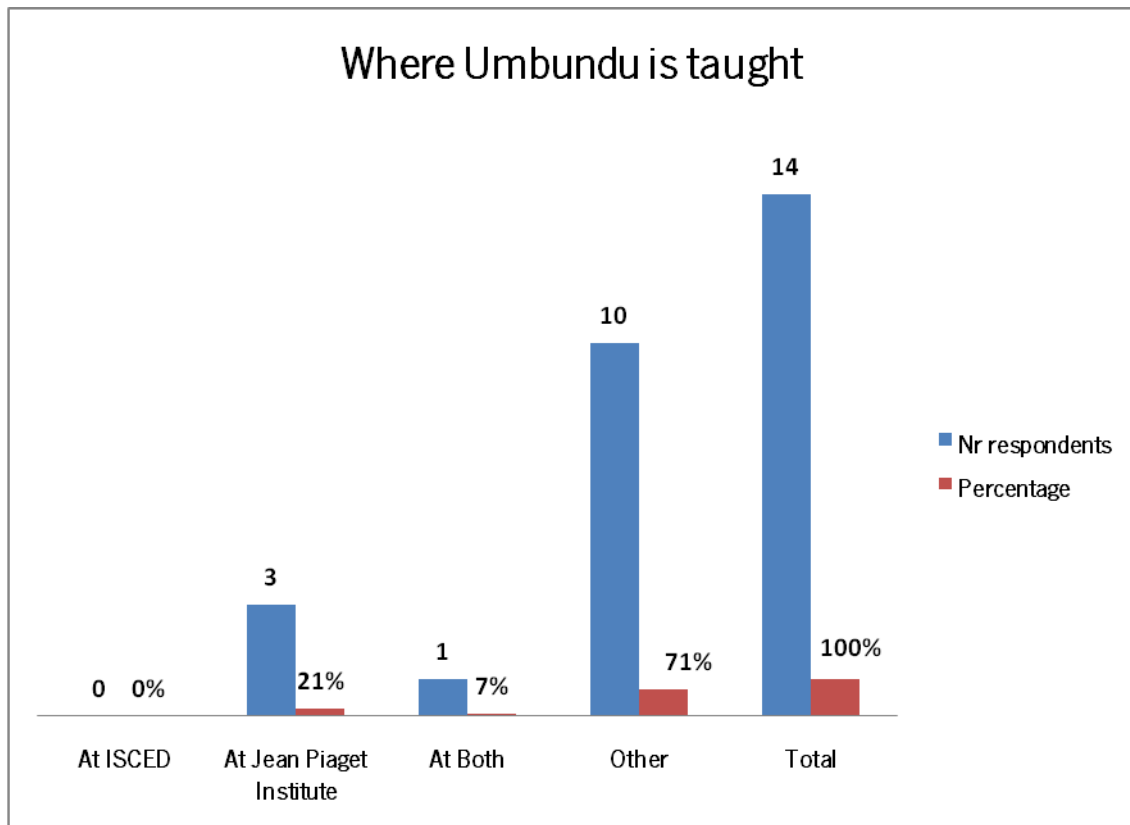
With respect to the learning of Umbundu, 93% of the respondents did not have it as a school subject (Fig. 31) and the percentage of those who can write and read Umbundu is relatively low (29%) but bigger than the other categories. This is an important element counting for the transference of linguistic knowledge in the family environment (Fig. 33). Fortunately, however, most of the respondents came to know Umbundu either at home or in the community and only 14% have had classes in Umbundu. In short, there cannot be enough academic background with relation to the Umbundu language as the most important part of the knowledge that there is has been transmitted

orally. The lack of enough literacy in the Umbundu language as provided by public schools hinders the awareness of the structure of the language at the orthographical level.

4.3.2.6. Group three: practice of Umbundu

The tenth question is “Where do you teach Umbundu? (Please, circle one option) (Onde é que ensinas o Umbundu? Por favor, circule uma opção)”

Figure 34: Places where the respondents teach Umbundu

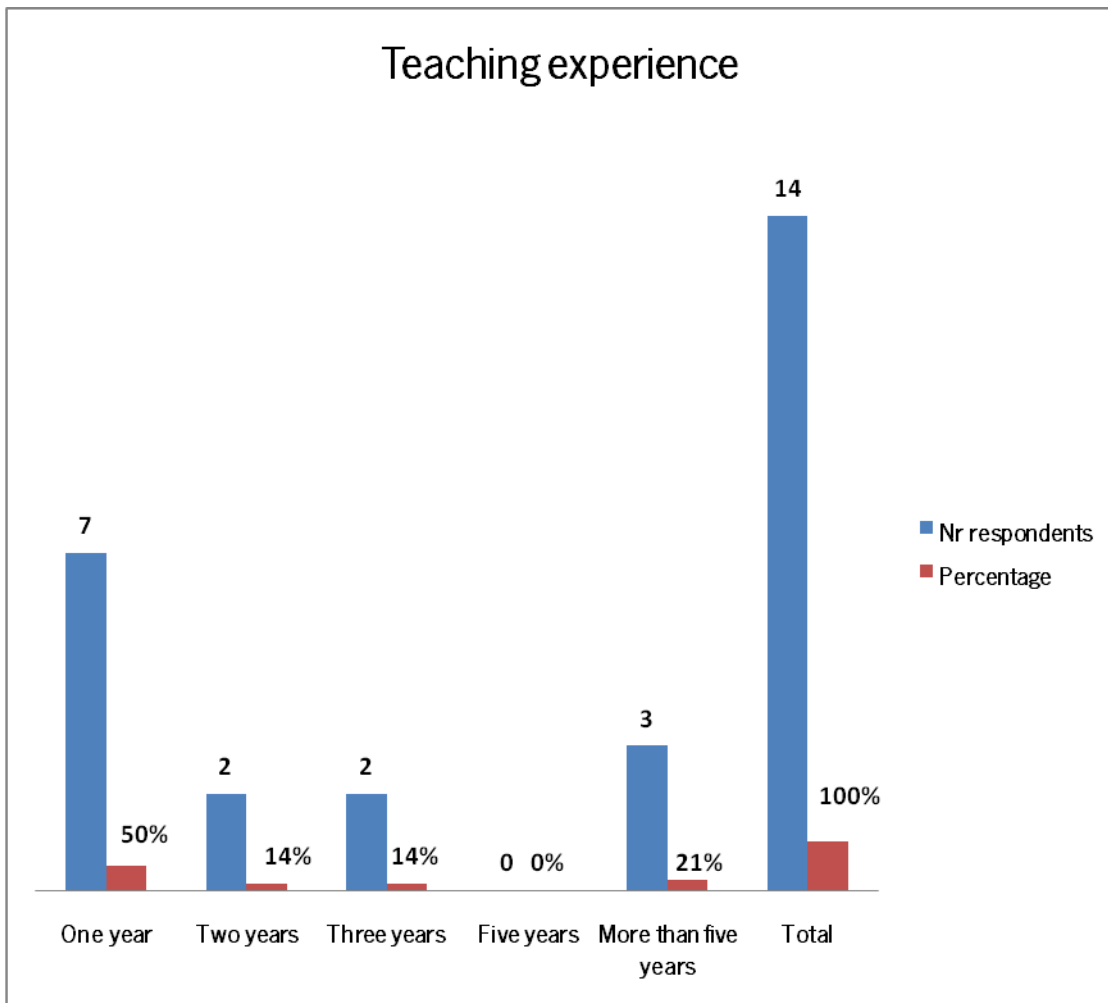


Note. This information is crucial as it tell about the institutions that give/teach Umbundu as a subject.

Figure 34 is about the places where Umbundu is taught. The blue bars shows the number of respondents and the red ones show the respective percentage. The graph reads that 21% of the respondents teach Umbundu at the “Jean Piaget” Institute and 7% indicate that they teach it at Piaget and ISCED (Instituto Superior de Ciências da Educação)/Benguela. 71% of the respondents indicate that Umbundu is taught in other places.

Question 11 reads like this “How long have you been teaching Umbundu? (Please, circle one option). (*Há quanto tempo ensinas o Umbundu? (Por favor, circule uma opção)*)

Figure 35: Respondents' time of teaching experience

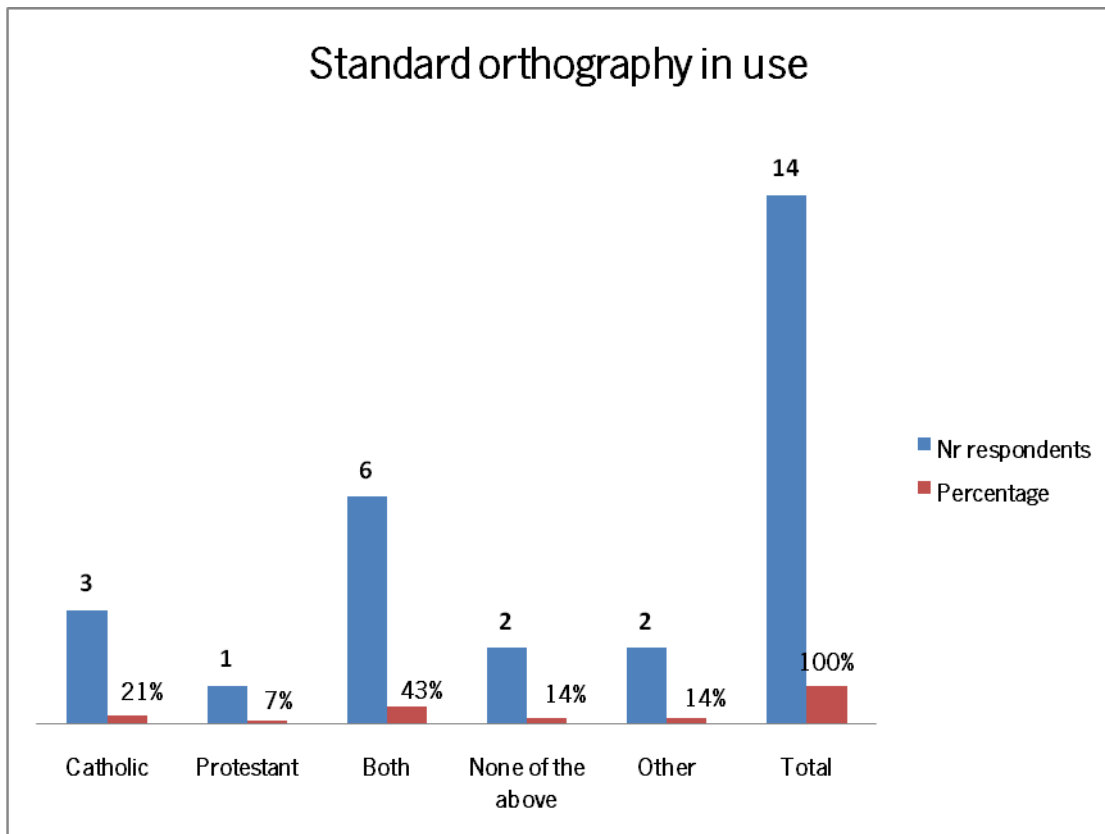


Note. Years of teaching experience can give an insight of the awareness teachers have of the teaching process itself, including the advantages and the drawbacks.

Figure 35 shows the respondents' teaching experience. The blue bars represent the number of respondents while the red bars represent the percentages. As seen in the graph 50% of the respondents have been teaching for a year and 14 % have been teaching for 2 years a similar percentage is attributed to those who have been teaching Umbundu for 3 years. Also, 21% have been teaching for more than 5 years.

Question 12 is written in the following way: "What Umbundu standard orthography do you use?) (Please, circle one option). (*Que padrão de ortografia da língua Umbundu é que usas? (Por favor, circule uma opção).*)"

Figure 36: The standard orthography each teacher uses

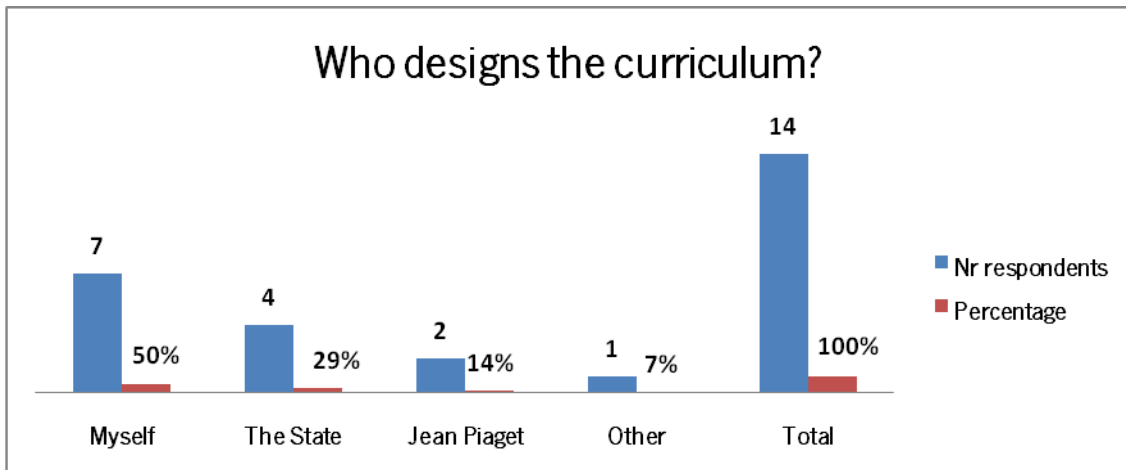


Note. The standard orthography that the respondents State they use for teaching can be taken to categorise the different approaches available to the writing of Umbundu.

Figure 36 shows the standard orthography that respondents use for teaching. The blue bars show the quantity of respondents and the red bars show the percentages. As it is displayed, 21% of the respondents use the Catholic orthography and 7% uses the Protestant Umbundu standard. Also 43% of the respondents use both the Catholic and the Protestant one.

Question 13 is “Who designs the curriculum of the Umbundu language teaching? (Please, circle one option). (Quem desenvolve o currículo para o ensino do Umbundo? (Por favor, circule uma opção) ”.

Figure 37: Who designs the curriculum?

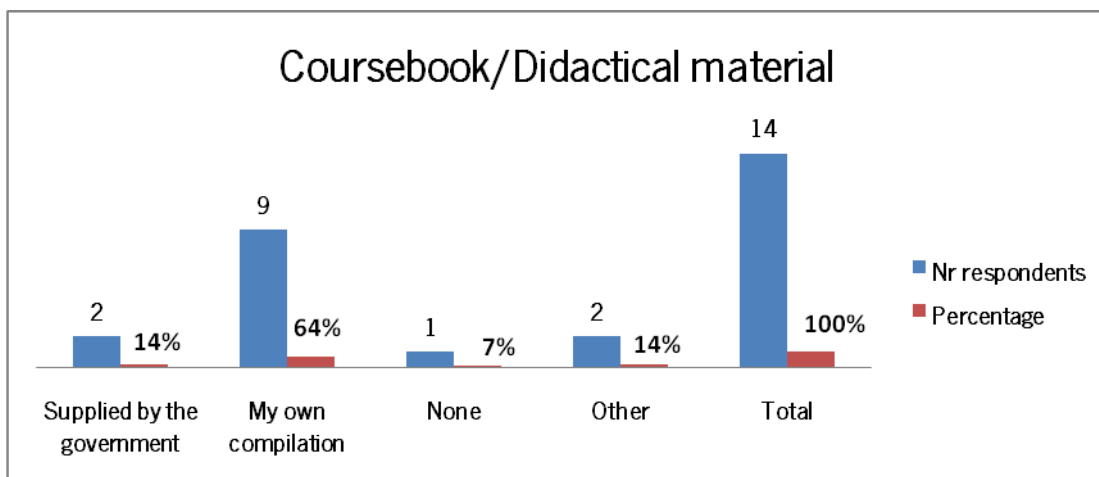


Note. The design of the curriculum here must be understood as the preparation of the programmes, the scheme of works and the didactic material for learners.

Figure 37 shows the results about who designs the curriculum, including the programme of Umbundu. The blue bars stand for the number of respondents while the red bars stand for the percentages. As it can be apprehended, 50% of the respondents design the curriculum themselves, 29% receive the curriculum from the State. Those who work for the “Jean Piaget” Institute receive the curriculum from there (14%).

Question 14 is “Which coursebook(s)/didactical material do you use for classes? (Please, circle one option). (Qual o manual/material didáctico que utiliza para as aulas?) (Por favor, circule uma opção) ”

Figure 38: Types of didactical material in use for Umbundu



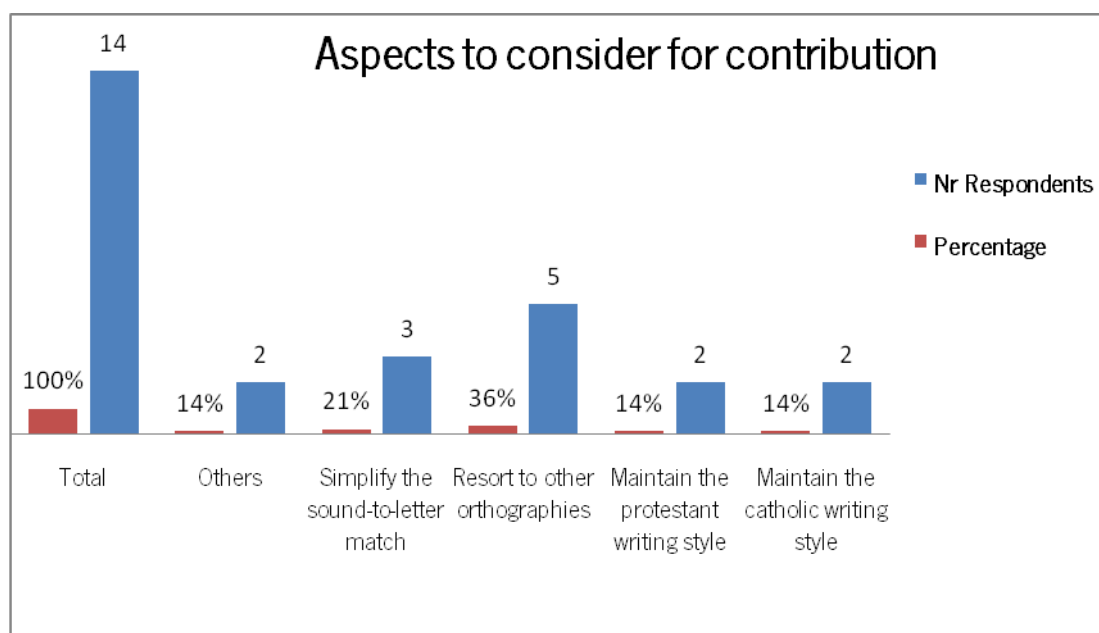
Note. Coursebook/didactical material may be produced by teachers or supplied by the government, including leaflets, grammars, some biblical texts and dictionaries.

Figure 38 presents the results about the supply of didactical materia for teaching. In the graph, the blue bars stand for the number of respondents while the red bars stand for the percentage.

As seen from the graph, the majority of the respondents (64%) compile their own coursebooks while 14% receive programmes from the State. 7% of the teachers has no coursebook supply at all.

Question 16 is “If you were asked to contribute to the harmonization of the Umbundu orthography for a common *corpus* planning for the education system, what of the aspects bellow would you like to help work out? (Please, circle one option). (*Se tivesses que contribuir para a harmonização da ortografia do Umbundu por uma planificação de corpus para o sistema de educação, que aspectos abaixo gostarias de te dedicar? (Por favor, circule)*”

Figure 39: Additional aspects that respondents consider for contribution on Umbundu orthography



Note. Aspects considered for contribution on the Umbundu orthography is an additional opportunity for the respondents to contribute about the Umbundu orthography towards its orthography.

Figure 39 displays the aspects respondents find important for contributing to the amelioration of the Umbundu orthography harmonization. As it can be seen, the majority of the respondents (36%) prefer resorting to other orthographic systems to ameliorate the existing ones. 14% had better maintain the Catholic standard orthography and a similar percentage had rather maintain the Protestant one. 21% prefer the simplification of the orthography to capture the sound-letter combination.

4.3.2.7. Conclusive remarks on the trainers’/teachers’ practice of Umbundu

Figure 34 shows that nowadays, Umbundu is taught by the respondents at both Piaget Higher Polytechnic Institute and ISCED – the Higher Institute of Educational Sciences in Benguela city. The teaching of Umbundu occurs in some of the institutions where the respondents work, being a sign

that its maintenance is being taken into due consideration. It is also an indication that teacher training may be taking place in the two institutions of higher education. So, it can be predicted that a ground is being paved for the preparation of people with sufficient professional knowledge to carry out the job of teaching the Umbundu language and reduce the existence of amateurs in public schools.

The majority of the respondents (71%) teach Umbundu at public schools in the phase of the experimental level (Fig.34). It is satisfactory that the majority of the respondents work at public schools. This means that the government has been engaged in the implementation of the Umbundu and other African languages in schools and material conditions are being created to support such an endeavor. It is important that at the experimental phase of the implementation of Umbundu teaching, the process is under the responsibility of teachers in State schools because it may prepare other State institutions (i.e. court, hospitals, the parliament and private schools) to get ready to respond to the demands of the spread of the use of African languages.

In terms of time of experience, the majority of the respondents (50%) have been teaching Umbundu for one year only and 21% have been teaching for more than five years (Fig. 35). The time teachers have had with the Umbundu language imparting inform that the generalisation of the teaching of African languages in public schools should already be a fact, give that some of the respondents are trainers and have been in this process for more than five years now. In fact the generalisation of the African language teaching can take place at any time depending on important factors explained under the principles of language planning discussed at the beginning of section 2.1.2.4., above, mainly the involvement of the principal decision makers and the language related internal coherence. At times, the generalisation of a project like this may depend more on the decision makers than the employed workforce.

Moreover, there is an indication of the respondents to mix various standards for the Umbundu language (43% of them); 21% use the CSO and 7% choose the PSO (Fig. 36). The respondents are divided with regard to the choice of the standard orthography they use, although it is clear that the majority use a mixture of both the CSO and the PSO. This may mean that there still are tendencies to teach Umbundu according to orthographical tradition of each respondent. This way, the teaching of the Umbundu language will risk being taught according to each teacher's likes of orthography (i.e. there will be a random choice of the orthography for the students). This is discouraging, mainly for learners (who may have different teachers in different years using different orthographical standards), teachers and material designers (who may be confused as to which standard to follow for the public).

As to the programme design for the Umbundu subject, 50% of the teachers do it by themselves, 29% of them receive the programmes they use from the State and 14% have programmes designed by Jean Piaget Institute (Fig. 37). If programmes for the Umbundu language are elaborated by different sources, the probability of the language to use many orthographical standards is high. Each school may decide to value the contents and principles enshrined in the documents they produce. This may already explain the existence of different orthographical standards seen in the tests in the appendices 7, 8 and 9 in comparison to the ones in the appendices 11, 12 and 13. This difference may be the reflection of the differing orthographical styles in the books that they use as classroom resource.

While 64% of them compile didactical material on their own, only 14% of the teachers have didactical material from the State, and the remaining 21% is of those who have not mentioned any specific material for the classroom (Fig. 38). The material for the classroom is more effective if it is issued by the same source. If teachers are in the position to compile their own materials for the classroom, the chance for them to confuse students of the language is high. Moreover, this is an indication that there is an urgent need to put discipline on which orthographical standard should be used and what it should look like.

In general, teachers tend to defend the orthographical standard of their own religion (making 14% for the Catholic orthography and 14 for the Protestant one). However, 36% believe that the Umbundu orthography should be reformed to meet nowadays needs of the education system and 21% of the respondents indicate that such reforms should result in a more simplified orthography (sound-to-letter correspondence) (Fig. 39). Here, the focus is more on the orthographical alternatives as seen by the respondents. Interestingly enough, it becomes more than clear that each teacher tends to use the orthograph of the denomination they belong to. The majority, however think that the orthography needs reform to make the orthography simpler than it is today. Others go farther by indicating that the reform should involve innovation as a result of taking advantage of other orthographical systems different from the Portuguese one on which Umbundu is currently written.

To conclude, it stands clear that the Umbundu language has been introduced into the education system at least at an experimental level and that this represent a fresh phenomenon in the education system. The tendency to mix different orthographical standards is an indication that it is urgent to start reforms of the orthography towards its harmonization for the language classroom. Moreover, didactical material has been designed by teachers themselves which reinforces the idea that there is not a harmonized orthography of the Umbundu language for the education system.

Finally, the need for a reform, despite the fact that there is the tendency of both Catholic Church members and Protestant members trying to support the maintenance of their own religion standard which may drive to conflicting desiderata between the two influential groups.

Given the results from both the questionnaires, presented in form of written samples and graphs, it stands clear that there has been enough evidence of the lack of harmonization with regards to the orthography for the Umbundu language as used by people from different religious backgrounds. As such, this research proposes practical ways towards the harmonization of this language's writing system, so that it can be used for the education system in a uniform way in the didactical materials, independent of the different writing traditions of teachers, learners and material designers. For the proposal to be sufficiently easy to follow, the current State of the orthography of the language will be presented in different figures and tables in the following chapter.

5. A PROPOSAL OF THE HARMONIZATION OF THE HUMBUNDU ORTHOGRAPHY

The antecedent chapter dealt with the information collected by means of questionnaires in both qualitative and quantitative forms. It considered the description of the respondents, the sample and the sampling techniques. The current chapter, however, presents a proposal of harmonization of the Umbundu orthographies at the level of the alphabet for the education system in Angola. For its concretization, the chapter has been divided into nine main subsections. At this stage now, it is important to quickly revisit the concept of harmonization, and then give a historical summary of how Umbundu started to be written. Thirdly, look at the current state of the Umbundu orthography through an analytical view of the Umbundu orthographic structures proposed by different authors and reveal its implications in the education system. The fourth subsection presents the translations made by the respondents to show how aware they are of the Umbundu orthography and how much they differ in terms of writing standards both among members of the same church and between members of different churches. The fifth subsection gives a short conclusion on the points discussed up to the third subsection. The sixth subsection looks at the current approaches to the Umbundu orthography in use by different institutions. The sixth section proposes what is believed to be a simpler, practical and updated harmonization of the orthography of the Umbundu language, at the level of the alphabet, by resorting to letters from the various orthographical systems existing in the world nowadays, as much as possible, to comply with the principles of “economy” (Batibo, 2000: 154) and that of convenience, that of conformity and that of consistency (Longtau, 2014). To understand the reasons for this proposal, it will be crucial to start with a description of a chronological development of the Umbundu orthography and its key players. The eighth subsection proposes the use of diacritics to differentiate homographs. The ninth subsection presents a proposal of the biblical passage of Mathew’s chapter 13, verses 1 to 9 described above.

5.1. The harmonization of orthography in subsaharan Africa: A conceptual framework

This section establishes the concept of orthographic harmonization, mainly as conceived in the context of orthographic studies in some countries of southern Africa. The importance of this concept in the context of this research resides in the fact that it has been construed differently by different authors. Babane and Chauke (2015) debate the harmonization of Xitsonga in South Africa and Shangani in Zimbabwe as trans-national languages which stemmed from the same parent language.

The political conditionalisms discussed in chapter 1 have made these languages to be written in unnecessary differing systems. These researchers are convinced that the orthographies of the two languages can be harmonized and believe that harmonization is a process of making “A single variety common to all the languages [...] be developed as a way of facilitating a closer unity among these languages spoken in different countries” (Babane & Chauke, 2015, p. 436). In their turn, Asher and Simpson (1994) refer to harmonization as a process happening when two or more dissimilar languages or dialects are fused to form one language free of contradictory features. In fact, the justification for languages to be harmonized lies in the fact that they are related, and that harmonizing them would make the written system more economical for users. Moreover, harmonization relates to the changes in the writing system of same or closely intelligible languages, so that these changes can facilitate linguistic interchange of its users by means of eye (and tactile) reading-based data. Examples of intelligibility (or its absence) between Xitsonga and Shangani below indicate the weight of orthographic standards in the direction that living languages of communities may take in the course of time. That is, the imposition of orthographical standards of European languages has potential to distort the nature of native languages, if serious interdisciplinary corpus planning of the languages involved is not taken into consideration and language-related experts are not involved. The way the languages in the comparative table of the figure below should be taken as a sign of how wrongly designed orthography may change the original languages as uttered by the native communities, during the learning and teaching processes. There are cases of distortion of community languages as a result of the imposition of European languages’ alphabets by non-native speakers (Jimbi & Sicala, 2020, p. 128)²⁸

²⁸ The aforementioned study by Jimbi and Sicala (2020, p.128) indicates:

- a) Some consequences of the imposition of the European alphabets (here, Portuguese) over the Umbundu language (i) confusion in pronouncing the letter <c> when followed by “a, o e u”. Cokwe can rigorously read [ˈkɔkwɛ] instead of [ˈtʃɔkwɛ] (ii) the confusion caused by the letter <s> when between <a> and <o>, making Sasoma be read [sa ˈzɔma] instead of [sa ˈsɔma] (iii) Double orthography, e.g., Kuando Kubangu, Kwanza-Sul vs Cuando Cubango, Cuanza-Sul (iv) the unnecessary use of trigraphs, digraphs and inappropriate diacritics, e.g., <Ng’> (PEDRO et al., 2013), <ñg> (CHIMBINDA, 2015) and <ñ> (DANIEL, 2002) for the phoneme [ɲ], when years before Valente (1964) suggested the use of the most economical <g̃>. (v) the imposition of the sound [r], resulting in “Chongoroi” (toponym) instead of “Tchongoloi” (vi) lack of consistency, e.g., “Umbundu” vs “Umbundo” (cf. QUINTA, BRÁS & GONÇALVES, 2017, p. 149), etc.

Figure 40: Xitsonga-Shangani's similarities testifying their relatedness

Verbs	
<i>Xitsonga</i>	<i>Shangani</i>
Sweka	sveka (to cook)
Raha	rhaha (to kick)
Vabya	vabwa (to get sick)
Veleka	b'eleka (to give birth)
Nouns	
<i>Xitsonga</i>	<i>Shangani</i>
Ririmi	lirhimi (language)
Xilo	cilo/chilo (a thing)
Byanyi	bwanyi/bwasi (grass)
Byalwa/Byala	bwala (beer)
Byongo	bongo (brain)
Dyambu	dambo (sun)
Swakudya	swakuda (food)

Note: This data can be found in Babane & Chauke,(2015, p. 438)

Figure 40 shows a comparative table of the relationship between Xitsonga and Shangani. The table of the figure presents the relatedness of verbs and nouns in the two languages. As can be seen, the slight differences between them can be accountable on the grounds of sovereignty. That is, language planning decisions made in one country was independent from the one made in the other. As a result, for example, the verb "to cook" is "*sweka*" in Shitsonga while it is spelt "*sveka*" in Shangani.

Another interesting study on harmonization was carried out by Kioko, Njoroge and Kuria (2012) where they discussed the unification of two Kenyan languages, namely the Gĩkũyũ and Kĩkamba. They consider two logics of orthographic harmonization (Kioko, Njoroge & Kuria, 2012, p. 56). On the one hand, there is a sense of deciding on which "graphemes and spelling" to take. On the other hand, there are decisions to be made on which writing system to prescribe if a language has not yet been given a writing system.

For the first case, it may be more difficult to negotiate the harmonization, as many agents may be tempted to defend their position, and so boycott the ideas of other stakeholders within this type context. For the second case, the facilitation resides in the fact that all harmonization-related issues may be negotiated in a common sensical way, as this becomes a case of orthography elaboration together. In orthographic harmonization, it is important to achieve common sense, so that the use of language for literary purpose adjusts to the best interest of the users and facilitate the writing of documents, the programming of technological devices such as computers and, more importantly, the teaching and learning processes of the language. Consistently, all inconsistencies should be

overcome to avoid the differences of writing the same sounds in different ways as shown in the table table of the figure below.

Figure 41: Orthography of the same sounds in the two Kenyan languages

	Speech sound	Gĩkũyũ grapheme	Kĩkamba grapheme
1.	/ɸ/	b	v
2.	/ʃ/	c	sy
3.	^h ɔ̃/	nj	ngy
4.	/tʃ/	c/ch	ky

Note: This table can be found in Kioko, Njoroge and Kuria (2012, p. 58)

Figure 41 shows orthography of the same sounds in the two Kenyan languages. In the table of that figure, it stands clear that the same sounds are written with different graphs in the two languages. As orthography affects the maintenance of the original aspects of languages, it can be that wrong decisions on the orthography cause wrong use of words as time goes by.

Harmonization of orthography needs to be carried out to the most minimal details possible. As such, all aspects of the languages to be harmonized have to be fully studied and compared, inconsistencies cleaned and consensus reached. That is, the grammatical aspects of the languages as well as its phonotactics should be re-elaborated towards the new orthographic conception for the language. Logically, the figure above shows two very different orthographies for what can be said the same language, a situation which may make the learning and teaching process difficult and the literary diffusion of knowledge by use of the language more complex.

The following is the analysis of the chronological evolution of the Umbundu language with focus on the aspects in which irregularities in terms of orthographic representation of sounds are more evident. A look at the incongruity in this regard may shed light on what made the Umbundu language to be written in so many different ways, which the agents have been and with what means they used to design the Umbundu orthography.

5.2. Chronological development of the Umbundu orthography and key authors

The section 1.3.1. reviewed the contributions made for the orthography of the Umbundu language by the most influential authors. It could be seen that the writing of Umbundu started, as it would happen to any casual encounter of cultures unknown to each other, with trying to ascertain what Umbundu as a language structure really was and how it could be translated into writing by means of very limited writing technology of the epoch. In this section a close look at how different authors created, adjusted and mastered the orthographical codification along the time is done.

Figure 42: Developmental traces of the Umbundu orthography

AUTHORS ↓	DEVELOPMENTAL TRACES OF THE UMBUNDU ORTHOGRAPHY (in 110-Year time)						
English words →	"hen"	"bat"	"rat"	"farms"	"dry season"	"small-pox"	"How much...?"
Koelle (1854)	"ösāndsī"	"erima"	"ómuku"	"owāpīa"	"ōkuēnye"	"odzingōngo"	"dşin.ōami...?"
Bleek (1869)	"osandyi"	"erima"	"omuku"	"ovāpya"	"okuenye"	"otyingongo"	"tyiñ.gami"
Capello & Ivens (1881)	"osandji" "osanji" "osan'qi"	"erima"	"omuçu"	"ovāpja"	"okuenhe"	"ot'chingongo" "otchingongo"	"T'chimgami?" "tchin'gami?"
Sanders and Fay (1885)	"osanje,"	"elima"	"omuku"	"ovapja"	"okwenye"	"Ocingongo"	"ciñame...?"
Torrend (1891)	"ossandyi," "osandyi"	"elima"	"omuku"	"ovapja"	"okuenye"	"Ocingongo"	"chingami...?"
do Nascimento (1894)	"osanji,"	"elima"	"omuku"	"ovapja"	"okuenhe",	"Ochingongo"	"chinhami...?"
Lecomte (1899)	"osangi,"	"elima"	"omuku"	"ovapja"	"okuenhe"	"Otchingongo"	"chiñgami...?"
Hambly (1934)	"osanji" "osandji"	"elima"	"omuku"	"ovapja"	"okweñe" "okwenye"	"Ocingongc" "otcingongc"	"chingami...?"
Alves (1964)	"osandji,"	"elima"	"omuku"	"ovapya"	"okweñe"	"Otjingongo"	"tjidami...?"
Valente (1964)	"osanji"	"elima"	"omuku"	"ovapya"	"okwenhe"	"Ochingongo"	"chidami"

Note. This figure shows a period of crucial 110 years of changes happening to 7 words and expressions of the Umbundu language as they were registered by different influential authors of their time.

The figure 42 gives a developmental spectrum of the Umbundu orthography as proposed by different agents with time. It has been revealed that Reverend Koelle's work, unless otherwise Stated, is the first written research that has ever existed on the Umbundu language. As a matter of fact, Koelle (1854, pp. 84 –120) describes Umbundu as a language of *Pangela*. Koelle was referring to a

language spoken in the Angolan region comprehending the Benguela province today. *Pangela* was the way he referred to the language whose name he could not really know from the natives. There was, however, a belief on his part that the region of the speakers was called that name. He believed that the language is also used in “Balundo, Viye, Síwūla, and Lubolo” (Koelle, 1854, p. 15). In his registrations, the Umbundu word for “tomorrow” is written “hēra” (op.cit., p.188), the word for “hen” is “ósāndši” (op.cit., p.127), the word for “bat” is “erima” (op.cit., p.125), the word for “rat” is “ómuku” (op.cit., p.123), the word for “farms” is “owāpīa”, the word for “dry season” is “ōkuénye”, (op.cit., p.91), the word for “small-pox” is “odzingōngo” (op.cit., p.55) As to the interrogative phrase “How much...?”, its Umbundu correspondent is written “dşin.ōami...?” (cf. *op.cit.*, p.99).

In the sequence of publications, there followed Dr. Bleek’s *Comparative Grammar of the South-African Languages*, published in London in 1869. In his orthographical codification, the word for “hen” is written “osandyi”²⁹ (cf. Bleek, 1869, p. 205ff), the word for “bat” is “erima” (cf. Bleek, op.cit., p. 205³⁰), the word for “rat” is “omuku” (cf. Bleek, op.cit., p. 205), the word for “farms” is “ovāpya”, the word for “dry season” is “okuenye”³¹, (op.cit., p. 205 - 207), the word for “small-pox” is “otyingongo”³² (op.cit. p.123), As to the interrogative phrase “How much...?”, its Umbundu correspondent is written “tyiń.gami” (cf. *op.cit.*, p. 219).

The third publication, in two volumes, was proposed by Capello and Ivens in 1881, explorers who influenced profoundly the way Umbundu and other Bantu languages of Angola came to be orthographically codified. In their orthographical codification, the word for “hen” is either “osandji” (cf. Capello & Ivens, 1881, v2, pp. 188 and 317), “osanji” (cf. Capello & Ivens, op.cit., v2, p. 159) or “osan’gi” (cf. Capello & Ivens, op.cit., v1, p. 226); the word for “bat” is “erima” (cf. Capello & Ivens, op.cit., v1, p. 130), the word for “rat” is “omuçu” (cf. Capello & Ivens, op.cit., v2, p. 130), the word for “farms” is “ovāpja” (cf. Capello & Ivens, op.cit., p. 43), the word for “dry season” is “okuenhe”, (cf. Capello & Ivens, op.cit., p. 355), the word for “small-pox” is either “ot’chिंगongo” or “otchingongo” (cf. Capello & Ivens, op.cit., pp. 159 and 161 respectively) As to the interrogative

²⁹ This spelling has been deducted from the way a similar sound has been represented in Bleek’s spelling. For example, “otyindyo” which mean “a big house”. This means that by that time the sound [dʒ] had changed from Koelle’s “-dş-” to Bleek’s “-dy-”. That means that rather than considering similar words in the references made, the focus should be on the way similar sounds are coded in not necessarily equal words.

³⁰ Here, the sound [r] still represents the Umbundu sound [l]. This may either result from mishearing the sound from the native speakers or by being influenced by previous literature, such as the one on the *Ojijherero* language. Bleek uses a similar orthography when refereeing to “eraka” (cf. Bleek, *op. cit.*, p. 205) which should be “elaka” in Umbundu (this language does not have the sound [r]). It is important that in research, researchers may easily get biased, mainly when there has been previous literature closely related to the object of study.

³¹ Here all the tonic accent have disappeared, but the sequence “-ny-” to represent the sound [ŋ] remains the same as used in Koelle’s above.

³² For this case, the use of “-ty-” substitutes Koelle’s “-dz-” (elsewhere “-dsi-”, [p. 29]) to represent the sound [tʃ].

phrase “How much...?”, its Umbundu correspondent is written “T’chingami?”/“tchin’gami?” (cf. Capello & Ivens, *op.cit.*, p. 79 and 113 respectively).

In 1885, Reverends Sanders and Fay wrote *Vocabulary of the Umbundu Language, Comprising Umbundu-English and English-English*. Their orthographical approach is that the Umbundu word for “hen” becomes “osanje,” (cf. Sanders & Fay, 1885, p. 17), the word for “bat” is “elima” (cf. Sanders & Fay, 1885, p. 5), the word for “rat” is “omuku” (cf. Sanders & Fay, 1885, p. 25), the word for “farms” is “ovapja” (cf. Sanders & Fay, 1885, p. 24), the word for “dry season” is “okwenye”, (cf. Sanders & Fay, 1885, p. 12), the word for “small-pox” is either “Ocingongo” (cf. Sanders & Fay, 1885, p. 50), the interrogative phrase “How much...?” is written “ciñame...?” (cf. Sanders & Fay, 1885, p. 17;).

Among the most influential contributions to the Umbundu orthography, it is worth mentioning Friar Jules Torrend of the Zambezi Mission. This missionary wrote *A Comparative Grammar of The South-African Bantu Languages Comprising Those of Zanzibar, Mozambique, The Zambezi, Kafirland, Benguela, Angola, The Congo, The Ogowe, The Cameroons, The Lake Region, etc.*, published in 1891 by the St-Austin's Press. In Fr. Torrend's orthography the Umbundu word for “hen” is “ossandyi,”/(also probably)“osandyi” (cf. Torrend's, 1891, p. 31), the word for “bat” is “elima” (cf. Torrend's, 1891, p. 31), the word for “rat” is “omuku” (cf. Torrend's, 1891, p. 31), the word for “farms” is “ovapja” (cf. Torrend's, 1891, p. 28), the word for “dry season” is “okuenye”, (cf. Torrend's, 1891, p. 31), the word for “small-pox” is “Ocingongo” (cf. Torrend's, 1891, p. 30), the interrogative phrase “How much...?” is written “chingami...?” (cf. Torrend's, 1891, p. 7).

José Pereira do Nascimento can be said to be the first native Portuguese who showed interest in the Umbundu language by evidentiating it with his *Grammatica do Umbundu ou Língua de Benguella*, in 1894. His orthographical approach is that the word for “hen” is written “osanji,” (cf. do Nascimento, 1894, p. 15), the word for “bat” is “elima” (cf. do Nascimento, 1894, p. 9), the word for “rat” is “omuku” (cf. do Nascimento, 1894, p. 21), the word for “farms” is “ovapja” (cf. do Nascimento, 1894, p. 13), the word for “dry season” is “okuenhe”, (cf. do Nascimento, 1894, p. 12), the word for “small-pox” is either “Ochingongo” (cf. do Nascimento, 1894, p. 38), the interrogative phrase “How much...?” is written “chinhami...?” (cf. do Nascimento, 1894, p. 98)³³.

³³In do Nascimento (1889, p. 98) the sequence “-nh-” in the underlined words of the sentence “Ombela ndanho i-loka-loka, onjo l'ovaue kai-fi, l'oku-vola kai-vola, i-kala-kala nho”, represents the sound [ŋ]. Again, in do Nascimento (*op.cit.*; p. 11) the same sequence represents the sound [ɲ]. This recursion to the same sequence to represent two different sounds may have resulted out of the lack of sufficient information about the Umbundu phonetics and phonology.

The most influential contribution for the orthographical design of the Umbundu language, from the Catholic church, today is the work of father Ernesto Lecomte, a priest of the Spiritan Order who wrote *Ondaka ia suku ou Doutrina Christa em umbundu e portuguez* (translated as *The word of God or Christian Doctrine in Umbundu and Portuguese*), published in 1899. His, orthographical codification of the Umbundu language, in most cases, has been used as a referential to the writings of ecclesiastic texts in the Catholic Standard Umbundu today. Taking the examples we have been presenting, the word for “hen” is written as “osangi³⁴,” (cf. Lecomte, 1899, pp. 30-33), the word for “bat” is “elima” (cf. Lecomte, 1899, pp. 30-33), the word for “rat” is “omuku” (cf. Lecomte, 1899, pp. 30-33), the word for “farms” is “ovapiã” (cf. Lecomte, 1899, pp. 30-33), the word for “dry season” is “okueñhe”, (cf. Lecomte, 1899, pp. 30-33), the word for “small-pox” is either “Otchingongo” (cf. Lecomte, 1899, pp. 30-33), the interrogative phrase “How much...?” is written “chiñgami...?” (cf. Lecomte, 1899, pp. 30-33).

In 1934, an anthropo-ethnologist called Wilfrid D. Hambly, Assistant Curator of African Ethnology, Field Museum of Natural History, published *The Ovimbundu of Angola* in which he includes a section on language. In his way of writing, the word for “hen” is “osanji/osandji³⁵,” (cf. Hambly, 1934, pp. 237, 249), the word for “bat” is “elima” (cf. Hambly, 1934, pp. 248), the word for “rat” is “omuku” (cf. Hambly, *op.cit.*, pp. 251), the word for “farms” is “ovapiã” (cf. Hambly, *op.cit.*, p. 246), the word for “dry season” is “okweñe”/“okwenye”/“okweñye”, (cf. Hambly, *op.cit.*, pp. 242, 237), the word for “small-pox” is “Ocingongo”/“otcingongo”/ (cf. Hambly, *op.cit.*, pp. 237, 245 respectively), the interrogative phrase “How much...?” is written “chiñgami...?” (cf. Hambly, *op.cit.*, p. 249).

Other two influential Catholic Standard Umbundu are found in the dictionary of Father Albino Alves entitled *Dicionário Etimológico Bundo-Português*, published in 1951, and in the grammar elaborated by Father José Francisco Valente entitled *Gramática Umbundu: A Língua do Centro de Angola*, published in 1964. Although the two are Catholic priests and writing about the same language, the way they approach the orthography for the Umbundu language shows important discrepancies. Taking the sample words being used for illustration, the Umbundu word for “hen” is

³⁴Lecomte (1899, pp. 30-33) codes the sound [dʒ] using the sequence “-gi” as seen in underlined words in “Ndati...va vayongia [,va'jondʒa] omanu kosiahulu...”, and in “...lu uongiolele [wondʒolele], lu uongiau[wondʒau]...”. It is worth noting here that Lecomte did not take into account the use of “j” to represent this sound which his preceding Umbundu writers used.

³⁵Hambly is not clear whether writing the sound [dʒ] with “j” or with “dj”. See for example “ohonji” on page 237 and “Kangándji” on page 249.

written “osandji,” with Father Albino Alves (cf. Alves, 1951, I, pp. 151) while with Father José Francisco Valente it is written “osANJI” (cf. Valente, 1964, p. 118); the Umbundu word for “bat” is written “elima” both in Alves (1951) and in Valente (1964). As to the Umbundu word for “rat”, “omuku” is taken by both the former and the latter. Moreover, the Umbundu word for “farms” is “ovapya” both in Alves (1951) and José Valente (cf. Valente, 1964, p. 31). The Umbundu word for “dry season” is “okweñe” in Alves (cf. Alves, 1951, V2, p. 985) and “okwenhe” in José Valente (cf. Valente, 1964, p. 161). The Umbundu word for “small-pox” is “Otjingongo” for Albino Alves (cf. Alves, 1951, V2, p. 1162), whereas in José Valente it is written “Ochingongo” (Valente, 1964, p. 31). As to the interrogative phrase “How much...?” its Umbundu correspondent is written “tjigami...?” in Albino Alves (cf. Alves, 1951, V2, pp. 1003 – 1004 & 1162) while in José Valente it is “chigami” (cf. Valente, 1964, pp. 25, 26).

The above chronological description represents the various ways the Umbundu orthography was proposed by different authors in the period ranging from 1854 to 1964, each one bringing to its writing rich influences from previous research realities exogenous to Africa, in general, and into the Ovimbundu linguistic region in particular. The ways the 7 illustrative lexemes in figure 45 above change with time and from one author to another may indicate that there was very little intergenerational orthographical communication, making their research more individualistic with tremendous consequences on how the language should be coded to respond to the principle of consistence and simplicity discussed earlier in Chapter 3. Unfortunately, many of those who attempted to write the Umbundu language do not seem to have made it with the intention of writing its orthography. Rather, they used the orthography of their own working languages, be they English, French, Spanish or Portuguese and adapted them to the Umbundu sounds. It means that the same word, as it appears in the figure above, might have been written according to the orthographical rules of the person who registered the linguistic events and items they came across. So, some ways are more approximate than the others. The use of <c> and <tch> <ch> <tsh> and <tj> to fill for the sound [tʃ] is an example of the influence of different orthographical rules of those who coded the African languages. One could predict that <c> is written for [tʃ] by a person whose education involves Romance languages such as Italian, <tch> by a person whose background is English and <tsh> for someone whose schooling is based on French. So, the Umbundu word [tʃove] – “yours” appears as “cove” in the PSO, “tchove” in the CSO and “tshove” by missionaries influenced by French.

For the following subsection, a focus is on the “Analytical Table of the Diachronic Evolution of the Umbundu Orthography”, where different ways of writing the same sounds have been collected from various contemporary literature.

5.3. Analytical tables of the evolution of the Umbundu orthography

The results from the Analytical table of the diachronic evolution of the Umbundu orthography (also available in a landscape perspective, in the Appendix 29), respond to the need to withdraw important elements of the journey that the Umbundu orthographic systems have done along the time. The table in the aforementioned Appendix has 7 columns and 29 rows. The 7 columns have been headed with the following categories: Sound (related to the phonetic symbols), example of corpus/source (related to the actual words in Umbundu as translated into English). Most probable influencing orthographical tradition (related to the probable orthographical roots which influenced the current ways of writing the Umbundu language), Adopting Religious Tradition (to verify which religious group has adopted which standard orthography), Author and publication (related to the authors of the sources considered in the examples, as well as the date of publications), Author’s Academic Status (it is a category that tells what specialization area of the author of a certain source is).

Figure 43: The sounds [ʃ] and [g]

Sound	Example of corpus/source	Most probable influencing orthographical tradition	Adopting Religious Tradition		Author and publication date	Author's Academic Status
			Catholic	Protestant		
[ʃ]	Oxako (leaves)	Portuguese, e.g.: Oxalá!	No	No	Chimbinda (2015: 30 – 31)	Catholic priest
	Isja (father of)	Umbundu, e.g. Isietu (from “ise” – father and “yetu” - our)	Yes	Yes	De Oliveira (2012: 922 – 923)	Portuguese Researcher/Teacher
[g]	Songela (sharpen)	English, e.g.: get, give. Kiswahili, e.g.: wageni; giza	Yes	Yes	SUZA, (2013: v) Palomé (1967: 39 and 42)	Coursebook designer Center for Applied Linguistics
	<i>ong'eleya</i> (grave)	Kokota language, Santa Isabel Island, e.g.: gato-õonu	No	No	Jimbi & Sicala (2020: 130)	Researchers
	Katengue (Toponym)	Portuguese, e.g.: formigueiro	No	No	Queiroz (1978: 112)	Journalist

Note. Analytical table of the diachronic evolution of the Umbundu orthography (author's data) – to be continued.

Figure 43 shows the sounds [ʃ] and [g]. As it can be seen, the two sounds are written in different forms by different authors. The first sound [ʃ] has been coded with a <x> by Chimbinda, a Catholic priest, although this codification has not been adopted by the CSO. In fact, both the CSO and the PSO have adopted <si> to code the sound [ʃ]. As to the second sound, [g], 3 ways of writing it can be identified. The use of <g> has been adopted by both the CSO and the PSO, although two other ways have been found (the use of <ḡ> and <gue>).

Figure 44: The sounds [ɲ] and [ŋ]

Sound	Example of corpus/source	Most probable influencing orthographical tradition	Adopting Religious Tradition		Author and publication date	Author's Academic Status
			Catholic	Protestant		
[ɲ]	ovo-nha (feather)	Portuguese, e.g.: minha, vinho	Yes	No	Do Nascimento (1894: 2)	
	Onyohā (snake)	Kiswahili, e.g.: nywi	No	Yes	SUZA, 2013: 18	Teachers
	K'eñulu (in the nose)	Spanish, e.g.: Spaña	No	No	De Oliveira (2012:1029)	Researcher
[ŋ]	Waniga (in feces)	Umbundu, e.g. Owigi	No	No	De Oliveira (2021 1027)	Researcher
	Ñgo (sole)	Umbundu, e.g. viñgami	Yes	No	Antônio, Cuianda & Bonifácio (2012: 10)	Course designers
	Ng'o (sole)	Umbundu, e.g.: ong'oma	No	No	Pedro <i>et al.</i> (2013: 14)	Researchers
	Osuñila (is dozing)	Umbundu, e.g.: Ñala	No	Yes	IESA (2009: 217)	Clergy

Note. Analytical table of the diachronic evolution of the Umbundu orthography (author's data) – to be continued

Figure 44 shows the sounds [ɲ] and [ŋ]. The sounds above are some of the most difficult to code in Umbundu. Their representation does not seem to find definitive response of the Portuguese alphabet. For example, the sound [ɲ] is written in 3 different ways (that is <nh>, <ny> and <ñ>) while the sound [ŋ] is written in 4 ways (that is, the use of <ḡ>, <ñg>, <ng'>, <ñ>). While the CSO takes <ñg>, the PSO and the JWSO take <ñ>.

Figure 45: The sound [tʃ]

punos	Example of corpus/source	Most probable influencing orthographical tradition	Adopting Religious Tradition		Author and publication date	Author's Academic Status
			Catholic	Protestant		
[tʃ]	Ochikumbu, (mutilated) Ochili, (it's true)	English, e.g.: much, chamber	No	No	Do Nascimento (1894: 2)	Researcher
	Tjituñu (Ravine) Tjimbundu (anthroponym)	Otjiherero, e.g.: <i>Otjiingirisa</i>	No	No	The New Otjiherero Dictionary English Herero Otjiherero Otjiingirisa (2010) available at https://dicg5fhppqu3ab.cloudfront.net/gkzjbmsnhulvfwfprxeqditoaac8c9bd3b8e6.pdf	A student of science
	Hatchoko (It's not)	English, e.g.: watch	Yes	No	Chimbinda (2015: 30) Malumbu (2007)	Catholic Priest, researcher
	Ucito (birth)	Italian, e.g.: vicino	No	Yes	Mezzadri (2003: 60)	Course designer

Note. Analytical table of the diachronic evolution of the Umbundu orthography (author's data) – to be continued.

Figure 45 displays the sound [tʃ]. The sound [tʃ] is another example of the difficulties that teachers and students can have in reading or writing in Umbundu. There are 4 ways of writing the same sound by different authors (that is <ch>, <tj>, <tch> and <c>). Moreover, there is no agreement between PSO and CSO as to writing this sound. While the PSO uses <c>, the CSO adopts <tch>. Different authors would write according to their conveniences (some write the sound with <tj>). Following the table of the figure, it can be seen that various authors have used unequal orthographies for the Umbundu words which are pronounced in the same way. It is this kind of differentiation that needs to be taken into account if the Umbundu language is to be implemented in the schools.

Figure 46: The sound [dʒ]

punos	Example of corpus/source	Most probable influencing orthographical tradition	Adopting Religious Tradition		Author and publication date	Author's Academic Status
			Catholic	Protestant		
[dʒ]	Onjevo (hunting)	English, e.g.: jack, judge	No	Yes	Do Nascimento (1894: 2) Hambly (1934: 237)	Medical Doctor African ethnologist
	Ondjali (progenitor)	English, e.g.: adjective	Yes	No	Chimbinda (2015: 30) Malumbu (2007)	Catholic Priest, Researcher

Note. Analytical table of the diachronic evolution of the Umbundu orthography (author's data) – to be continued.

Figure 46 shows the sound [ɖʒ]. Another relevant sound is [ɖʒ] which appears in two main different forms. One adopted by the CSO (<dj>) and another more used for the PSO (<j>). These two standards have been adopted by the majority of the users. The question now is which one is to be implemented with du justification

Figure 47: The sounds [s]; [k]; [w]

Sound	Example of corpus/source	Most probable influencing orthographical tradition	Adopting Religious Tradition		Author and publication date	Author's Academic Status
			Catholic	Protestant		
[s]	Oḡoma (King)	Kiswahili, e.g.: sambusa, sisi	Yes	Yes	SUZA (2013: v)	Course designers
	Saçoma (anthroponym)	Portuguese, e.g.: moça	No	No	Jimbi & Sicala (2020: 130)	Researchers
[k]	Ualungu ka (smart)	Kiswahili, e.g.: Karibuni,	Yes	Yes	SUZA, 2013: 16	Course designers
	Çunene (big)	Portuguese, e.g.: Cabinda	No	No	Queiroz (1978: 93)	Journalist
	Caluqu embe (Toponym)	Portuguese, e.g.: quando	No	No	Queiroz (1978: 93)	Journalist
[w]	Umue (somebody)	Portuguese, e.g.: nua, atuar	No	Yes	Do Nascimento (1894: 6)	
	Wanda (has gone)	English, e.g.: watch	Yes	Yes	Chimbinda (2015: 88)	Catholic priest, researcher

Note. Analytical table of the diachronic evolution of the Umbundu orthography (author's data) – to be continued.

Figure 47 displays the sounds [s]; [k]; [w]. The sound [s] has been used for the PSO, CSO, JWSO and the SSO. An alternative way of writing this sound by resorting to the use of <ç> has been proposed by Jimbi & Sicala (2020, p. 130). As to the sound [k], 3 alternatives have been evident in texts. One can find <k>, <c> and <qu>. The CSO, the PSO, the JWSO and the SSO have adopted <k> for this sound. One may easily find <c> and <qu>, mainly with regards to the writing of proper names. Finally, the sound [w] has been used in two forms. While the PSO and the JWSO adopt <u>, the CSO and the SSO take <W> to write the sound.

Figure 48: The sound [j]

Sound	Example of corpus/source	Most probable influencing orthographical tradition	Adopting Religious Tradition		Author and publication date	Author's Academic Status
			Catholic	Protestant		
[j]	lene (yours)	Portuguese, higiene	No	Yes	Do Nascimento (1894: 7)	
	Yapa (meanwhile)	English, e.g.: Yes	Yes	Yes	Chimbinda (2015: 88)	Catholic priest, researcher

Note. Analytical table of the diachronic evolution of the Umbundu orthography (author's data) – to be continued.

Figure 48 represents the sound [j]. Another sound that is represented in different ways of writing is [j]. The figure above shows that it is written either as <i> or as <y>. While the CSO uses <y>, the PSO uses <i>. Also, different authors take different approaches to write it, depending on whether they have had a Catholic or a Protestant education.

The above figures show the various ways that some sounds of the Umbundu language are written. It stands clear that different institutions and agents adopt different ways of writing the same sounds, making it difficult to choose the one to use, mainly at the State schools where different teachers and learners may come from different religious education which, in the end, have significant influence in the way the textual contents are orthographically presented. The following subsections will deal with the transcribed samples of the translation made by users of Umbundu who participated in the study.

5.4. Practice of Umbundu 3 (translations into Umbundu by respondents)

There has been a necessity to ask respondents for translations into Umbundu. The question 22 goes as follows: “Try the most to write the following sentences in Umbundu, so it can be inferred how skillful you are in it: (*Tenta o teu máximo de escrever as seguintes frases na língua Umbundu para que se saiba o quanto podes*)”

Respondents have been asked to translate sentences (see translations attempts in better visible picture in the Appendix 21, in a landscape perspective) with the objective of, first, seeing how far they are aware of Umbundu as language of their community, the value that is expected to be given by its users, what they used it for, with what frequency and how they care for it in everyday life; secondly, tracing some of the written patterns mirrored in their translations and compare them to both the Catholic and Protestant orthographies, with focus on letter combinations representing specific sounds in Umbundu.

Figure 49: Variation of the Umbundu orthography in translations according to religion

Words	Umbundu pronunciation [Phonetic transcription]	Catholic respondents	Number of varying forms	Protestant (Evangelic) respondents	Number of varying forms
Baby (noun)	[ɔŋaŋa]	ngāngā; o ngannga; nganga; o ngānga; oŋgaŋga	5	oñaña; o ñaña; nānā; onānhā	4
Lord (noun)	[ŋala]	Ngāla; ngala; ngala	3	Ñala; nala; nālā	3
Witch (noun)	[ɔnganga]	Onganga; o nganga	2	Onganga	0
Sprinkled (verb)	[wanana]	Wanhānhā; wanhanha; wanyanya	3	wañaña; wanhanha; wanhānhā; wanyanya	4
Uses to defecate (phrasal verb)	[jinjanja]	Inyanya; inhanha; iniania	3	Iniania; iniāniā; yinyanya	3
Slithers (verb)	[lifelena/lisjelena]	Lilisielenena; lisyelena; lichelena; liatchelena	4	Lisielena; lichelena	2
Yawn (noun)	[ɔnana]	Onhanha; onyanya	2	Onhanha; Onhānhā; onyanya; onyānyā	4
Misery (noun)	[ɔngongo]	Ongongo; o ngongo	2	Ongongo	0
Hunchback (noun)	[ɔŋɔŋɔ]	Ongōngō; Ongōngō; ongongo; ongongo	4	Oñoño; onōnō	2

Note. A table showing the variations in the translations' samples done by the respondents

Figure 49 shows the *corpus* selection, representing solely a small portion of the intricacies that challenge the orthographical arrangements of the Umbundu language. Not only discrepancies can be seen between religious writing traditions, but also the constant irregularities within the same religious group. The number of variations tends to be as much as the number of respondents. Only rarely can one see consistence within the Protestant responses as seen with the words “onganga” and “ongongo”. These evidences are faithful testimony of the Umbundu language background the respondents represent and the literary intensity with which they use the Umbundu language.

5.5. Conclusive remarks on practice of Umbundu 3 (translations into Umbundu)

The translations were asked for the purpose of evaluating how far the respondents are aware of the orthographical rules of the Umbundu language. The table that shows the variations in the translations made by both Catholic and Protestant members is a demonstration of the current instability of the Umbundu orthography. For the first sound pair [ŋ]/[g], the Catholic members have varied 6 times to represent [ŋ] while the Protestants varied 4 times. On the other hand, the Catholic members varied twice while the Protestant had only one version. For the second sound pair [ɲ]/[nj], the Catholic members vary 4 times representing the sound [ɲ] while the Protestant ones only vary 5 times. For

the sound [nj], the Catholic members showed 3 variations while their Protestant counterpart showed 4 versions. For the sound [ŋ], the Catholic members varied 5 times while the Protestants varied 4 times. In the long run, the evidences are crystal clear that the variations show the instability of the orthography of the Umbundu language as learnt and applied by both Catholic members and Protestant ones, and that there is a clear sign of weak literacy in the Umbundu orthography.

5.6. Current orthographical approaches to Umbundu

The current State of the Umbundu orthography is characterised by the existence of various ways of writing. As it will be seen, the excerpts below have been extracted from the Umbundu Bibles as written by the Catholics (the excerpt on the left) and the Protestants (the excerpt in the middle left), the Jehovah Witnesses (the excerpt in the middle right) and the government (the excerpt on the right).

Teachers and learners usually debate the authorized standard for the education system. However, the chaos continues in the samples of the texts on which they base their way of writing. The Catholic, the Protestant and the Jehovah Witnesses' versions of the Umbundu orthography still mismatch in many instances and so does the State's.

Figure 50: The four ways of the Umbundu orthography per standards



Note. The figure shows the CSO, the PSO, the JWSO and the SSO respectively. The colours show same sounds written in relatively different ways. The texts have been rewritten in the following table so that it can be better read.

Figure 50 shows biblical passages in Mathew, chapter 13, from verses 1 to 9, as written by the Catholic (Catholic Standard Orthography or CSO), the Protestant (Protestant Standard Orthography or PSO) and the Jehovah Witnesses (Jehovah Witnesses Standard Orthography or JWSO) respectively, and a sample text from an alphabetisation manual from the Angolan State (State Standard Orthography or SSO).

Transpositions of the CSO, the PSO and the JWSO have been made in table 2 below. The way they have been written reflect, to a large extent, the chaotic traces of the orthography of the Umbundu language in the appendices 19, 20 and 21 as written by different authors and the respondents of this research, respectively. In fact, the information of the chaotic and not uniform way of the orthography of Umbundu pervades the test designs taking place in various different schools in Benguela province (See Appendices 6 – 18).

Table 2: Transposition of the text of Matthew Chapter 13, verse 1 to 9

Transposition of the CSO	Transposition of the PSO	Transposition of the JWSO
<p>1V'eteke lyatcho, Yesu watunda v'imbo kwendje wakatumala k'ekulo lya kalunga. ² Nahō awiñgi alwa alitutila kokwahe, kwendje oñgwala mo: omo lyatcho, eye wañgila v'owato omo atumala osimbu awiñgi atalama k'ongongo ya kalunga. ³ Yesu yapa wavalombolwila ovina vyalwa v'alusapo. Tete wavatila olusapo olu, heti:</p> <p>⁴Ngundja umwe waenda k'epya lyahe okuwaya ombuto. Etchi akala l'okuwaya, olomema vimwe vyalokila k'onele k'onele y'elila. Olondjila vyeya kwendje vyavilya. ⁵ Vimwe olomema vikwavo vyawila v'otchikalambwe, ndakuti, v'otchikakala, muna mwakala lika eve litito; olomema vyatcho evi vyatunda l'ombili omo kuti osi kayalongele. ⁶ Pwāyi utanha etchi wakatwa, vyovuka kwendje vyakukuta, omo olombombo kavyañgilile p'osi. ⁷ Olomema vikuavo handi vyalokila v'olosongo kwendje vyakulila kumosi lavyo; pwāyi noke olosongo vyeya okuvitotomisa. ⁸ Olomema vyakamukwavo vyawila v'osi yiwa; ovyo kwendje vyama epako: lumwe otchita, l'ukwavo akwi epandu, l'ukwavo vali akwi atatu. ⁹ Ū okwete atwi okuyeva, ayevelele!</p>	<p>1Keteke liaco, Yesu wa tunda konjo kuenje wa tumāla kohulo yokalunga. ² Kuenje owiñi walua wa liongoluilu kokuaye, yu wa londa vowato umue, wa tumāla kuenda owiñi wosi wa talama kohulo yokalunga. ³ Noke wa va lombolwila ovina vialua pocakati calusapo, loku popia hati: "Tala! Umue ukuakulima wa tunda oku waya ombuto. ⁴ Osimbu a kala oku waya, olombuto vimue via kupukila konele yonjila kuenje olonjila vieya via vi lia. ⁵ Olombuto vikuavo via kupukila pocikalakala, pana okuti ka pa kaile osi yalua kuenda via kula lonjanga momo osi yaco ka ya longele. ⁶ Pole, eci utanya wa tua, via kakuka kuenda viowuka, momo ka via kuatele olombombo. ⁷ Vikuavo via kupukila pocambanda kuenje eci ocambanda ca kula, ca totomisa ovikūla viaco. ⁸ Handi vikuavo, via kupukila posi yiwa kuenje via fetika oku ima apako, vimue ci pitahāla 100 kolonjanja okuti, evi via kūliwa vi sule yikuavo 60, handi yikuavo 30. ⁹ Una o kuete ovatūi, a yevelele."</p>	<p>1Keteke liaco Yesu wa tunda vonjo, yu wa tumala kohulo yokalunga. ²Owiñi wa lua wo liongoluilu oco wa wila vowato, kuenje wa tumalamo, owiñi wosi wa kala kongongo. ³Yu wa popia lavo ovina via lua lolosapo hati, Uwayi umue wa tunda oku waya. ⁴Eci a kala loku waya, olomema vimue via lokila kohulo yonjila kuenje olonjila vieya via vi lia. ⁵Vikuavo via lokila pocikalakala, apa ka pali eve lia lua, oco via tunda lonjanga momo eve ka lia luile. ⁶Eci ekumbi lia tunda via kakuka, kuenje, omo ka via kaile lolombombo, viovuka. ⁷Vikuavo via lokila pocambanda kuenje ocambanda ca yova kuenje ca vi totomisa. ⁸Vikuavo via lokila posi yiwa kuenje via ima olomema, vimue ocita, vimue akui epandu, vimue akui atatu. ⁹U o kuete ovatui oku yeva a yeve.</p>

Note. Matthew Chapter 13, verse 1 to 9 rewritten for better comprehension³⁶.

³⁶ **13** That same day Jesus went out of the house and sat by the lake. *Such large crowds gathered around him that he got into a boat and sat in it, while all the people stood on the shore. *Then he told them many things in parables, saying: "A farmer went out to sow his seed. *As he was scattering the seed, some fell along the path, and the birds came and ate it up. *Some fell on rocky places, where it did not have much soil. It sprang

The table 2 is the transposition of the original texts from the Bible above (Fig 50). They are all written in Umbundu. The texts have been marked with 4 colours, in the figure above, used to underline the words and mark the differences in the orthographical patterns that they bring. The first colour is the red one used to mark the difference between two words which are pronounced in the same way in Umbundu and mean the same thing but written in different ways. For example, one can have “lyatcho” with the CSO and “liaco” with both the PSO and the JWSO and adopted for the SSO (as shown in the right column).

The following tables show a selection of words to illustrate the differences as they appear according to the standard orthography. The SSO has been intentionally brought here solely for the sake of illustration of the way of writing the similar sounds.

Table 3: Synopsis of how the sound [tʃ] is written according to each standard

THE SOUND [tʃ] IN 4 WRITTEN APPROACHES (SIMILARITIES AND DIFFERENCES)			
CSO <tch>	PSO <c>	JWSO <c>	SSO<c>
Ly <u>atcho</u> (line 1)	Li <u>aco</u> (line 1)	Li <u>aco</u> (line 1)	<u>Ca</u> lwa (line 2)
Et <u>chi</u> (line 13)	E <u>ci</u> (line 7)	E <u>ci</u> (line 18)	<u>O</u> cili (line 9)

Note. Sounds are written in different ways according to religion.

Table 3 shows a synopsis of how the sound [tʃ] is written according to each standard. The table shows examples of 8 words which have the sound [tʃ] written in the 4 standard orthographies. It can be seen that the Catholic orthography is more dissimilar to the others.

The second colour, in Figure 50, to consider here is the blue one used to show that while the CSO is consistently conjunctive, the PSO (Protestant Standard Orthography) tend to be more disjunctive and the JWSO is like the Protestant one but with the quantifiers “walua” and “vialua” (here in pink colour in the column on the middle right) are written conjunctively. The State’s Umbundu writing approach currently used is more similar to the JWSO.

The discrepancies above show the lack of consensus as to which orthography is to be followed. This explains the existence of possibilities to write Umbundu by resorting to any form that can convince the writer or reader of being right without the need to account for it.

up quickly, because the soil was shallow. *But when the sun came up, the plants were scorched, and they withered because they had no root. †Other seed fell among thorns, which grew up and choked the plants. *Still other seed fell on good soil, where it produced a crop—a hundred, sixty or thirty times what was sown. *Whoever has ears, let them hear.”

Table 4: Synopsis of how the conjunctive and disjunctive approaches differ per standard

Conjunctive vs the disjunctive approaches			
CSO (Conjunctive)	PSO (disjunctive)	JWSO (conjunctive)	SSO(conjunctive)
Alwa (line 4)	Wa lua (line 3)	Walua (line 4)	Calwa (line 1)
Vyalwa (line 10)	Lia lua (line 10)	Vialua (line 8)	Vyaco (line 6)

Note. The differences in writing the same expressions of Umbundu according to religion, shpwing the preference either of the disjunctive or the conjunctive approach to the Umbundu orthography.

Table 4 shows Synopsis of how the conjunctive and disjunctive approaches differ per standard. The lack of consistence about whether to predict one approach or another is clear. Besides, there is a clear indecision about the use of characters to represent sounds. For example, “Wa lua” in line 3 of the PSO differs from “Walua” in line 4 of the JWSO.

The third colour, in Figure 50, to consider is the green one. It is used to establish the difference between the two orthographies in coding the velar sound [ŋ]. While in the CSO the velar sound is coded as “ñg”, in the PSO and in the JWSO it is coded as “ñ” in the Umbundu word for “crowd” one finds “owiñgi” (CSO) and “owiñi” (SPO and JWSO). The State manuals write this sound in a way that is similar to that of the CSO.

Table 5: Synopsis of how the sound[ŋ] is written per standard

THE SOUND [ŋ] IN 4 WRITTEN APPROACHES (SIMILARITIES AND DIFFERENCES			
CSO <ñg>	PSO <ñ>	JWSO <ñ>	SSO< ñg>
Awiñgi (line 3)	Owiñi (line 4)	Owiñi (line 3)	Kowiñgi (line 5)

Table 5 represents a synopsis of how the sound[ŋ] is written per standard. At least, here, a marked difference is seen with the CSO in the word “Awiñgi” (line 3 of Figure 50) and the SSO in the word “Kowiñgi” (line 5 of Figure 50) compared to their counterpart in the PSO and in the JWSO.

There also is a difference in how nasal vowels are signaled by a tilde in both the JWSO and the SSO and not in the CSO or in the PSO. For example, the Umbundu words for “ears” is written “atwĩ” in the CSO, “atui” in the PSO “atũ” in the JWSO. Looking at the way the word for “tenth” is written (“yakwĩ”) one may well deduct that the SSO would write it in the way the CSO does.

Table 6: Synopsis of how nasalized vowels are accentuated

HOW NASALIZED VOWELS ARE WRITTEN ACCORDING TO EACH STANDARD			
CSO <generally no diacritic>; rarely <î>	PSO <no diacritic>	JWSO <ã > <ũ>	SSO<ã> <ê>
wakatum <u>ala</u> (line 2) atw <u>ĩ</u> (line 32)	Wa tum <u>ala</u> (line 2) Ovat <u>uj</u> (line 18)	Wa tum <u>ãla</u> (line 2) Ovat <u>ũ</u> (line 28)	om <u>ãla</u> (line 2) amal <u>êhe</u> (line 9)

Table 6 is a Synopsis of how nasalized vowels are signaled. For example, in the phrase “Wa tumala” (PSO, line 2 in Table 6) the nasal sound is not signaled with a *tilde* diacritic, while “Wa tumãla” (JWSO, line 2 in the same table) is marked with a *tilde* diacritic.

The brown and the violet colours in Figure 50 are used to signal the difference between the way the sounds [dʒ] and [tʃ] are written in the CSO, in the PSO, in the JWSO and the SSO versions. For example, the [dʒ] sound in the Umbundu words for “and” which is “kwendje” (in the CSO) and “kuenje” (in both the PSO and JWSO) the difference is clear. The SSO version writes this word in a mixed way, i.e. “kwenje” (here the difference is in the use of the glide sound <w> in the CSO and the State manuals as opposed to the use of the <u> in both PSO and JWSO and the use of <dje> in the CSO as opposed to the use of <j> in both the PSO and the JWSO). For the [tʃ] sound, in words such as “etchi” and “otchili” (CSO) and “eci” and “ocili” (both in the PSO and the JWSO) the difference in the way they are written is clear. The SSO writes these words in a way similar to the PSO and JWSO.

Table 7: Synopsis of how [dʒ] is written per standard

HOW [dʒ] IS WRITTEN ACCORDING TO EACH STANDARD			
CSO <dj>	PSO <j>	JWSO <j>	SSO<j>
Kwendje (line 2)	kuenje (line 2)	kuenje (line 2)	Olonjanja (António, Cuianda & Bonifácio, 2012, p. 19)

Table 7 shows a synopsis of how the sound [dʒ] is written per standard. As can be seen, this sound is written in the same way in all the standards of Umbundu in the table but the Catholic one. For example, while the Catholic version is “Kwendje” in the first column of the table, the other versions write “kuenje”.

The differences in the way the four orthographical standards conflict indicates the existence of *a priori* circumstances and reasons for their conception. So, It can be coherently Stated that there are discrepancies in the way the Umbundu language is written by four strongly influential systems with clear direct consequences in the teaching, learning and standardization processes, essentially achievable by means of consensus by all the players and not by dictated “politically-driven” decisions.

The analysis of the results in the tables of the figures above mirror the influence of the political changes that occurred with the colonial division of African States, mainly with the role played by the Berlin Conference as discussed in Chapter 1. The Conference determined limites which made different linguistic communities – constituting real States/kingdoms – become part of new empire (be it Portuguese, English, French, German or Spanish) whose language police was oriented to the imposition of the language of the colonizer and the demise of the autochthonous languages, on the one hand. On the othe hand, the Missionary Services made it possible to give life to local languages because of their Christian message to the entire world. As the work was made differently in different colonial territories, some of the langauge related work got more advanced than the others. Umbundu languages studies started to be seriously touched by non-Portuguese speaking missionaries who had to write it according to the languages of their original empire. As left clear in the section 1.3., the Portuguese language policies were blunt enough to deter the teaching of the local languages in the missions under the famous decree 77, published in the Angola’s Official Bulletin (*Boletim Oficial de Angola, 1921*), Nr 5, 1st series (December 9, 1921). As consequence, Umbundu and other langauges were, for a greayer period of time taught in secrecy in Missionary schools. This made Umbundu be coded by two main schools: the Catholic Missions and the Protestant Missions. The main difference between the Catholic approach and the Protestant one was the way the former were more pro-empire and the latter more pro-local communities. This made the majority of the Umbundu community become more linked to the PSO than to the CSO. Besides, the cathecisms and the Bible translations into Umbundu were first made under the anglophone missionaries coming from Canada, the USA and England, or those from other countries who were in their missions. The use of Umbundu as means of instruction was not so much carried out in the Catholic institutions as it was in Protestant missions, let alone the use of Umbundu in the public schools. Using Umbundu was against the law and the Umbundu was considered a “language of dogs”. While the Protestant missions’ members could deal with the natives in their own languages, translated Christian songs

into Umbundu and made it valuable, the Catholic church members were oriented to speak Portuguese as much as necessary.

In fact, the first Catholic Bible in Umbundu was published in the post independence era. Its version was written without any consultation of their Protestant counterpart. The lack of communication between the two religious institutions made the emergency of the two main orthographical standards in force today. Sporadically, there are other standards developed both by individual agency – take the case of the researchers presented in chapter 1 and in chapters 4 and 5 – and cooperative agencies – take the case of JWSO and the SSO.

All in all, this section has focused on the presentation of the data collected by means of documents analysis. The results from the comparative analysis of the texts from the Bible (one written in the CSO, the PSO, the JWSO and excerpts of texts from the Manual of Umbundu (SSO). Mathew, 13 from verse 1 to 9 were analyzed and traces of existing discrepancies were highlighted to facilitate the identification of aspects of the orthographical standards informing the differences in writing. The following section, in its turn, presents a proposal of what is thought to be a simpler, more practical and updated harmonization of the Umbundu orthography towards its effective implementation in the education system.

5.7. A proposal of harmonization of the Umbundu orthography: Alphabet and word parts

In this section, a proposal of the harmonization of the Umbundu orthography will be first made, mainly at the level of the alphabet by adopting and adapting letters which either did not exist in the current orthography before or those letters which already exist. Secondly, the orthographic harmonization needs to be made, mainly at the level of morphosyntactic analysis. The table in the appendix 23 shows the sounds and the way they are written in the PSO, the CSO, the JWSO, SSO and other individually developed ways of writing (i.e., heterographies), followed by punctual remarks on the options for an ameliorated orthography for the classroom.

The New Proposal (NP)/Current Adoption (CA) concerns the offer of alternative alphabetic possibilities (adding to existing range of letters, new ones from non-Portuguese alphabets) and/or adjustments as found more feasible for teaching and learning. In fact, the table in the following subsection is an attempt to present a proposal of a harmonized orthography of the Umbundu language at the level of the alphabet, as perceived to conform to the principles of economy, convenience, conformity and consistency and make the orthography shallower. A consideration of the

use of newly introduced characters has been taken to fill in the gaps of the current alphabetical system of Portuguese in responding to phonotactics of Umbundu.

5.7.1. Previous attempts to ameliorate the Umbundu alphabet

The alphabet adopted by Malumbu (2007) and Valente (1964) is, basically, the Latin one made of 22 letters. In that case, the sequence would be as in the following table:

Table 8: The previously accepted alphabets

Majúscules	A B C D E F G \tilde{G} H I J K L M N O P S T U V W X Y ³⁷
Minuscules	a b c d e f g \tilde{g} h i k l m n o p s t u v w x y

Note. Malumbu (2007) does not take the letter < \tilde{g} > as he has decided for the digraph < ñg >.

Table 8 shows previously accepted alphabetic letters for the Umbundu language. The basis for this alphabet is the Latin one. In this alphabetic list some letters which the Portuguese alphabet use as special characters for foreign words (e.g. <K>; <Y>; <W>). In some works such as that of Valente (1964), < \tilde{g} > is used to represent the nasal velar sound [ŋ], another complex sound for speakers of Umbundu.

5.7.2. The recent proposal of harmonization of the alphabet for the Umbundu language

Attempts have been made to propose orthography for the Umbundu language with potential to respond to the phonotactics of the Umbundu language (e.g., Jimbi & Sicala, 2020; Pedro *et al.*, 2013; Malumbu, 2006; Valente, 1964). The table below is one of these proposals.

³⁷ The letters K, W and Y have been adopted from the Swahili alphabet influence (Valente, 1964, p. 24). As to Pedro *et al.* (2013), the most recent study of the Angolan languages' orthographic harmonization, for the Umbundu orthography, the sequence "mb", "nd""ng", "ng'" "nj" and "ny" are taken as elements of the alphabet, and this is proposed for the manuals in the education system. The question is that the alphabet is generally made of sequence of letters from A to Z and not expected to be made of sequence of diphthongs or triphthongs.

Table 9: The adopted alphabet proposal for the Umbundu language

Maiúsculas	A B Ç Ć D E F G ġ Ĝ H I J K L M N Ñ O P S Š T U V W X Y
Minúsculas	a b ç ć d e f g ġ ĝ h i j k l m n ñ o p s š t u v w x y

Note. This proposal can be found in Jimbi and Sicala (2020, p. 130)

Table 9 shows a recent proposal for the Umbundu language. The proposal of the alphabet above comes as an attempt to mitigate the current gap in terms of one-to-one correspondence between the segments of the spoken and the written Umbundu. One of the consequences of the application of this alphabet is that the character <S> will only be Used as initial for the sound [s], making the intermediary[s], be written with the letter <ç>. Additionally, <Ć/ć> is another innovation for the sound [tʃ] which ends with the confusion made by the use of the current alternatives. With it there will be no more need to use the digraph <ch>, the trigraph <tch> or the monograph <c>. There also is the introduction of the character <ġ> (macron) for the hard [g]. This character resolves the existing problem with the use of <g> and its behaviour when followed by <e> and <i>. In the same line is the character < ĵ >. This one has been proposed here to fill for the sound [dʒ]. Currently, there are two ways of writing it: < dj> (CSO) and <j> (PSO, JWSO and the SSO). <Ñ/ñ> which are currently used for [ŋ] with the PSO are proposed to function as they are used in Spanish and so replacing the use of the digraph <nh> (CSO) and <ny> (PSO). In that case, sound [ŋ] is proposed to be written using <ğ> which is a nasalized version of the non-nasalized hard [g]. Lastly, not the least, <Š/š> are suggested to play for the sound [ʃ], a rare sound in Umbundu which also constitute a difficulty for readers and writers.

All this innovation on the orthography for the Umbundu language results from the consciousness of the real need to offer an alternative that can not only work as a way of having the language recorded, but also, and most importantly, made with the purpose of preserving, as much as possible, the roots of the Umbundu language which risks alienating its original sounds, not only for the sake of natural course along the time, but more for the incapability of the European languages' alphabets to fully respond to this important coverage that all African languages spoken in Angola deserve.

At the morphosyntactic level, there are potential critical changes that can occur so as to facilitate the understanding of the structural components of meaning of words in Umbundu. As

such, some decisions need to be made to more accurately decide when and where the orthography should be conjunctive or disjunctive.

As to the approach to the Umbundu orthography, there are many problems related to whether some semantic unities should be written conjunctively or disjunctively. It is, therefore, necessary to harmonize these important issues as they have potentials to confuse learners at the level of contrastive analysis, mainly because their first literacy experience has been gained in Portuguese. (A few cases will be presented in the paragraphs that follow).

The basic criteria to be considered has been the analysis of the long neglected complex forms of the verb “to be” which tend to change according to the grammatical categories of the words they are attached to. So, the criteria will also include the separation of words according to the grammatical categories they really belong to instead of agglutinating the verb “to be” and the immediate elements of its phrasal structure. This way is believed to facilitate a contrastive analysis of structures and sentence elements with students in the classroom.

In general, there has been confusion when it comes to distinguish a form of the verb “to be” because in Umbundu it is agglutinated with the noun. Most learners may find it very difficult to discern these compounds as nouns as learners are, in the main, influenced by the disjunctive construct in Portuguese.

In Umbundu, the conjunctive approach to the morphosyntactic arrangements seems to be influenced mainly by the work of the philologist Wilhelm Heinrich Immanuel Bleek (and those who came after him), wherein the famous ‘classes of Bantu languages’ noun prefixes’ have been established (Bleek, 1869). For the *Nano* (the name he gives to the Umbundu language) 15 such classes of nominal prefixes are proposed and discussed in appealing details (Bleek, 1869, p. 217). Although these nominal classes have been increased to a number of 18 (Simões *et al.*, 2020, p. 4), there is a strong suspicion that instead of noun prefixes, they are pure verbal forms of the verb “to be” (“o kukala” in Umbundu) as it will be demonstrated in the following comparative tables with examples and proposals of harmonization towards disjunction, as we are, in fact, in presence of two free morphemes in the conjunctively arranged “classes of nominal prefixes”.

A very practical example of how Bleek has influenced the way scholars tend to approach structural analysis of the verb “to be” as prefix of nouns, adjectives and prepositions, just to mention some aspects, is found in Sasoma (2015, p. 38ff), a researcher who is a native user of Umbundu. It can be perceived, unless otherwise explained, Professor Sasoma seems to describe the “o-“, in the words “omola” (literally “ is child”) and “omunu” (literally “is human”) under the first class of such

“prefixes” corresponding to their “plurals” with the prefixes in class 9, so that one can have “omala” and “omanu”, respectively. What Professor Sasoma seems to take as plural prefix is “oma-” which substitutes the presumable singular prefix “O-“to make the plural forms. However, what may not have been understood, thus far, is that “o-” is not a nominal prefix. It is, instead, a free morpheme corresponding to the English “is” (third person of “to be”) which is linked to the noun “mola” which, in its turn, corresponds to the English “child”. This kind of nouns makes the plural form by substituting the vowel of the first syllable in “mola” with an infix (“-a-”), resulting in “mala” which corresponds to the English plural “children”. This can easily be proven true in the sentence “Uu o mōlā” (i.e. “this is child”) as contrasted with “Ava o mālā”(i.e. “these are children”). It sounds strange for any investigator based on Bleek’s nominal classes of prefixes. Here, “uu” fills for “this”, “o” fills for “is” and “mōlā” fills for “child”, so that in the plural form, the sentence becomes “ava” standing for “these” and “o” standing for “is” and, finally, one has “mālā” standing for “children”. In the same way, one has to approach the morphosyntactic analysis of “omunu” (standing for the literal “is human”) whereby one has “o” standing for “is” and “munu” standing for “single human”. So, the plural of “munu” must be, in rigorous terms, written as “mānu” (“various humans”). In the same way, the plural “mōla” is made by substituting the vowel of the first syllable by the infix “-ā-” which gives it “mālā” standing for “children. These are irregular (exceptional) passages from the singular to the plural.

Table 10: Demonstration of why a disjunctive orthography would be more functional for the learner

How it is currently written (cf Sasoma, 2015)	Demonstrative	Verbal form "to be"	Noun form	Harmonization proposal (disjunctive)
Uu "Omola" (Noun, singular) ava "omāla" (Noun, plural) ³⁸	Uu/ ava	"o"/ "o"	"mōla" "māla"	"uu o mōlā" (this is child) "ava o mālā" (these are children)
Uu "omala" Ava "vamala"	Uu Ava	"omala" vamālā	Non-existent Non-existent	Uu omālā (This finishes) Ava vamālā (these finish)
English literary translation	This These	is are	child children	This is child These are children

Note. With the disjunctive orthography, teaching and learning the Umbundu language becomes more facilitated.

Table 10 shows Demonstration of why a disjunctive orthography would be more functional for the learner. In fact, it is an exercise to demonstrate how the disjunctive approach to the Umbundu orthography helps understand the grammatical classes that may get opaque if used in a conjunctive way.

Another analysis can be done with regard to what Bleek calls the 2nd, the 3rd and the 4th classes of nominal prefixes ("u-"). In Sasoma (2015, p. 39), one finds "utungi" (is builder), "ungombo" (is herder) and "ulongisi" (is teacher) respectively. The plural for "u-", in these classes, is "'a-". In that case, one has the following plural phrases: "a tungi" ("a tunḡi") standing for "are builders", "a ngombo" standing for "are herders" and "a lonḡisi" standing for "are teachers", respectively.

³⁸ A case of irregular plural formation of the verb "to be". Similar cases are "omu nu" (is person) whose irregular plural form is "oma nu" ("are person").

Table 11: Demonstration of why a disjunctive orthography would be more functional for the learner 2

How it is currently written (see Sasoma, 2015)	Demonstrative	Verbal form “to be”	Noun form	Harmonization proposal (disjunctive)
Uu “utungi” (sing.) ava “atungi” (plu.)	Uu ava	“u” ”a”	“tungi”	“uu u tunḡi” “ava a tunḡi”
English literal translation	This These	is are	constructor constructors	This is constructor These are constructors

Note. This way, it becomes clearer for a student whose first language of instruction is Portuguese.

Table 11 displays another demonstration of why a disjunctive orthography would be more functional for the learner. The explanation given above is applicable to the understanding of this table. The proposal of how the orthography can be harmonized has been put at the last column.

The third analysis on Bleek’s influence can be considered with regard to his classes 8, 13 and 19 as discussed in Sasoma (2015, p. 39). Taken as prefix, “o-“ in “ongonga”, “ombwa” and “ofeka” respectively, would indicate the singular and the corresponding plural prefix would be “olo-“ of Bleek’s class 10 (Simões *et al*, 2020, p. 4) so that the results could be “ongonga” (hawk) and “olonganga” (hawks); “ombwa” (dog) and “olombwa” (dogs); “ofeka”(country) and “olofeka”(countries). Actually, those “prefixes” are real forms of the verb “to be” which would result in “o ngonga” (is hawk) vs “olo ngonga”(are hawk); “o mbwa”(is dog) vs “olo mbwa”(are dog); “o feka” (is country) vs “olo feka” (are country).

Table 12: Demonstration of why a disjunctive orthography would be more functional for the learner 3

How it is currently written (see Sasoma, 2015)	Demonstrative	Verbal form "to be"	Noun form	Harmonization proposal (disjunction)
Ei "ongonga" (sing) Evi "olongonga" (plu)	Ei evi	"o" "olo"	"ngonga"	"ei o ngonga" "evi olo ngonga"
English literal translation	This These	Is are	Hawk	This is hawk These are hawk

Note. This way, explaining the structures becomes facilitated.

Table 12 shows, once more, demonstration of why a disjunctive orthography would be more functional for the learner. Leaning on the analyses made above, one can easily find out that one is not in presence of nominal prefixes but free morphemes representing the verb "to be" in both the singular and the plural. As seen above, the nouns take no plural morphemes. The plural or singular are determined by the subject and reflected in the verb.

Breaking this barrier is not found easy for a researcher whose paradigm has been Bleek's. However, it is important to understand that elements which used to be called "nominal prefixes" are pure forms of the verb "to be" which depend much on the kind of nouns they link to.

Table 13: Splitting verb + noun agglutinations

How it is written (conjunctive)	Verbal form “to be”	Noun form	Harmonization proposal (disjunction)
O mbêmbwa (peace)	O (is)	Mbêmbwa (peace)	Eçi O Mbêmbwa (this is peace)
E p ata (family)	E (is)	Pata (family)	Eçi E pata (it is family)
A p ata (families)	A (are)	Pata (family)	Aa A pata (these are family)
Û k ãĩ (woman)	Û (is)	kãĩ (woman)	Uu Û kãĩ (this is woman)
A k ãĩ (Women)	A (are)	Kãĩ (woman)	Ava A kãĩ (these are woman)
O ng ombe (cow)	O (is)	Ngombe (cow)	Eçi O ngombe (this is cow)
O l ongombe (cows)	Olo (are)	Ngombe (cow)	Evi Olo ngombe (these are cow)
O ç ikomohiso (wonder)	Oçi (is)	Komohiso (wonder)	Eçi Oçi komõhĩço (this is wonder)
O v ikomoiso (wonders)	Ovi(are)	Komohiso (wonder)	Evi Ovi komõhĩço (these are wonder)

Note: Continued

Table 13 proposes a disjunctive paradigm for the arrangements of words. In this proposal, a separation of parts (disjunctive approach) wherein a clear separation between the conjugation of the verb “to be” and nouns is preferred (i.e. free morphemes have to be separated in the sentences). The position taken here is based on the fact that what has been called classes of noun prefixes so far (Simões *et al.*, 2020: 4 – 6) is, in fact, conjugations of the verb “to be”. The advantage of a disjunctive approach is that it facilitates a more comprehensive learning and teaching process of students whose first language learning experience is Portuguese, that is, different grammatical categories should be split.

Table 14: Splitting verb + adjective agglutinations

How it is written (conjunctive)	Verbal form “to be”	Adjective form	Harmonization proposal (disjunction)
Inene (big)	i (is)	nene (big)	Ei i nene (this is big)
Činene (big)	či (it is)	nene (big)	Eči či nene (this is big)
Vinene (big, plural)	vi (they are)	nene (big)	Evi vi nene (these are big)
Vanene (big, plural)	va (they are)	ene (big)	Ava va nene (they are big)
Ūçõvĩ (tall)	Ū (it/she/he is)	sõvĩ (tall)	Uu Ū sõvĩ (this is tall)
Açõvĩ (they are tall, plural)	A (they are)	sõvĩ (tall)	Aa A sõvĩ (these are tall)
Vaçõvĩ (they are tall, plural)	Va (they are)	sõvĩ (tall)	Ava Va sõvĩ (these are tall)

Note. Similar processes are found in this table.

Table 14 shows the way verb + adjective agglutinations have been split. Then, a critical description is needed here so that a conjugation of the verb “to be”, which used to be considered an affix (nominal prefix) in Bleek (1869, p. 217) and those who came after him such as Werner (1919, p. 38ff) and Schadeberg (1982, p. 111) is revealed. Taking this into consideration, the adjective “nene” in the table above, is an entirely free morpheme (cf. Simões *et al.*, 2020, p. 6) agglutinated with the conjugation of the verb “to be” (another free morpheme) in form of “i”. This revelation makes it possible to suggest a disjunctive approach to the morphosyntactic arrangement proposed for the harmonization of the Umbundu orthography. This way, the learning and teaching processes become more facilitated for the approach already exists in the learning and teaching of Portuguese.

Table 15: Splitting verb + preposition/adverbs: towards harmonization (disjunction)

How it is written (Conjunctively)	verbal form	preposition	Harmonization proposal (disjunctive)
İñgilamo/iñilāmo (he/she enters into)	iñgĩlā/iñila (he/she enters)	mo/mõ (into)	iğĩlā mō (enters into)
Wakalapo (he/she was there)	wakala (he/she was)	po (there/here)	wakala po (he/she was there)

Note. Similar processes taking place in the table above.

Table 15 shows the way verbs, prepositions and adverbs (free morphemes) are made one under the conjunctive approach to the Umbundu orthography. Following what has been explained about

the previous tables, it is plausible to propose a disjunctive approach to the “verbe+preposition” and “verb-adverb” order in Umbundu. The reasons are didactical ones as discussed above, which makes the teaching and learning facilitated.

Table 16: Splitting preposition + verb + noun agglutinations

How it is written	Preposition	Verb+noun agglutination	Harmonization proposal (disjunctive)
Vimbo (vo+i+mbo) (in what is the township)	vo (in the)	imbo (i+mbo) (is the township)	V'i mbo (in the township)
Komanu (on what is the people)	Ko (on the)	Omānu (oma+nu) (they are people)	K'o mānu (on the persons) Lit. On (those who) are persons

Note. In this table, agglutination of final vowels when in contact with beginning vowels takes place. In writing, the preposition and the verb are agglutinated.

The table 16 shows how preposition + verb + noun agglutinations can be split. That is, the way a conjunctive approach has made three grammatical classes as one, i.e., a preposition (“vo”), a conjugation of the verb “to be” in the third person singular (“i”) and the noun (“mbo”) are solely interpreted as a noun (“Vimbo”) and been taken as such. The conjugation may have been ignored for its apparently strange location in the sentence and as forced by what is seen as a long transmitted misunderstanding of the verb form of “be” in many works of researchers (see, for example, Schadeberg, 1990, p. 10; Simões *et al.*, 2020, p. 4ff, just to mention a few) who take the verb forms as classes of nominal prefixes.

5.8. The use of diacritics to differentiate homographs

Many words in Umbundu usually are written in the same way, although they have different meanings. For these words, diacritics as already suggested in Schadeberg, (1990) should be used over respective constituent vowels so as to facilitate both the teacher and the learner to help each other in the correct pronunciation of the words. The following table presents some of these homographs.

Table 17: The use of diacritics to differentiate homographs (disjunctive proposal)

How it is written	How it is written	Harmonization proposal	Harmonization proposal
Ulela (lubricant/oil)	Ulela (take care of her/him)	Ū léla (is lubricant/oil)	Ulèlà! (take care of her/him)
O kukuta (to wrap up)	O kukuta (he/she dries up)	O ku kúta (is the wrapping up)	Okukùta (he/she dries up)
Ohumba (it is basin)	Ohumba(it is disease)	O hũmba (it is basin)	O humba(is disease)
Okuvela (it is being sick)	Okuvela (to provoke)	O ku véla (is the sickness)	O kuvèla (to provoke)
Onete(it is chest)	Onete (it is mole)	O nété(is chest)	O nètè (it is mole)
Onjimbi (it is singer)	Onjimbi (it is Marabu Stork, a big night bird)	O n̄jimbi (it is singer)	O n̄jĩmbi (it is Marabu Stork)

Table 17 shows the use of diacritics to differentiate homographs (disjunctive proposal). There are various reasons why the orthography of the Umbundu language should consider the standardization of the use of *tilde* (˜) as a diacritic for the nasalization of graphemes. In fact, Schadeberg (1982, p. 109) has pointed out that Umbundu,

...provides unusual data on all aspects of nasals and nasalization: phonetic, phonological, and diachronic. These facts cannot be gathered from the existing descriptions of UMBundu which contain no more than scattered allusions to some of the nasalized segments. A recent publication which is concerned mainly with spelling suggestions for Angolan languages provides a good start for an inventory of nasalization in UMBundu but the conclusions reached are largely unfounded and even self-contradictory (*Instituto Nacional de Linguas, 1980*).

The study above has provided the 10 phonetic segments of the Umbundu language in a methodical way. The figure below shows how the nasalization of the majority of the letters in use should be done. Unlike the Portuguese orthography that only offers possibilities of nasalizations with the tild over <ã> and <õ>, Schadeberg (1982, p. 110) acknowledges the use of this diacritic over other characters, such as <v>; <l>; <h>; <y>; <w>; <i>; <e> and <u>. This acknowledgement guarantees that languages such as Umbundu can be better preserved by maintaining their original spoken versions as advocated in Jimbi and Sicala's "*Orthographic entrepreneurship: a proposal of harmonization and preservation of the Umbundu language spoken in the south-centre region of*

Angola”, published in the journal “Revista Digital de Políticas Lingüísticas” (Jimbi & Sicala, 2020, p. 128).

Figure 51: Schadeberg’s systematic phonetic segments of Umbundu

p	f	b̃	v	ṽ	m	y	ỹ	i	ĩ
t	s	d	l	l̃	n			e	ẽ
c		j			ɲ			a	ã
k		g			ŋ			o	õ
	h			h̃		w	w̃	u	ũ

Note. As in Schadeberg, (1982, p. 110)

Figure 51 displays Schadeberg’s systematic phonetic segments of Umbundu. Taking the segments above into due consideration, the difference between very similar and, thus, confusing ways of writing and reading Umbundu words can be clear to the learner. The reason for it resides in the fact that a person who knows Portuguese perceives the structures of words and may learn quicker through a comparative analysis of the two languages. To facilitate the learners’ assimilation of the Umbundu language, the way the sentence is arranged should be as closer to the language they already know. The following table shows some examples to illustrate how the use of the tilde makes a great difference.

Table 18: The use of tilde over letters to differentiate words

How it is written	How it is written	Harmonization proposal	Harmonization proposal
Walēla (he/she is fat)	walela(he/she is light)	Walēlā (he/she is fat)	Walèla (s/he is light/not heavy)
Olosi (lands)	Olōsi (they are little fishes)	Olo si (are land)	Olō sī (are small fishes)
Okutōla (it is being strong)	Okutola (it is tearing)	O ku tōlā (is strengthening)	O ku tola (is tearing)
Ovāla(they are guts)	ovala (he/she is difficult)	O vālā (are guts)	Ovala (hurts/is difficult)
Otela (he/she crafts) iron)	otēla (he/she knows)	Otela (crafts (iron))	Otēlā (knows)
Onya (he/she defecates)	Onya (it is envy)	Onya (defecates)	O ñā (is envy)

Table 18 shows the use of tilde over letters to differentiate words. As Umbundu is much nasalized, these segments need to be accurately signaled. This table presents an exercise on it. It can be seen that the nasalized segments have been made evident by resorting to the tilde diacritic.

The following subsection looks at the proposal of how the biblical passages in Umbundu should be written to help common users, learners and teachers better enterprete the countours of the parts of words and how they combine to make meaning.

5.9. A proposal of the biblical passage of Mathew’s chapter 13, verses 1 to 9

The following text is an exercise to demonstrate how the chapter of the bible shown above (taking the PSO and the JWSO examples) could be written in a way that Umbundu becomes easier to teach and learn leaning on a more disjunctive approach. This way has been proposed for its potential benefits in terms of making it easy for the learner to understand the disposition of the most important syntactic and morphological elements of the sentences.

«1Ke teke lia čo Yesu watunda vo njo, yuu watumālā ko hulo yo kalunga. 2 O wingĩ wa lwa wo liongolwila o čo wawila vo wato, kwenje watumālā mō, o wĩgi wo si wakala ko ngongo. 3 Yuu wapopya la vo ovi nā vya lwa l’olo sapo hati, U wayi umwē watunda o kuwaya. 4 Eči akala lo kuwaya, olo mēmā vi mwē vialokila ko hulo yo njila kwenje olo njila vieya viavilya. 5 Vi kwavo vialokila po či kala-kala, apa ka pali eve lya lwa, o čo viatunda lo njanga mómō eve

ka lialwile. 6 Eči e kumbi liatunda viakakuka, kwenje, ómō ka viakale l'olo mbombo, viovuka. 7 Vi ku avo vialokila po ča mbanda kwenje o ča mbanda čayova kuenje ča vi totomiča. 8 Vi ku avo vialokila po si i wa kwenje viaimā olo mēmā, vi mwē o čita, vi mwē a kuī e pandu, vi mwē akuī a tatu. 9 Uu o kwete ova tuī o kuyeva ayeve.»³⁹

The following is a proposal for the SSO as presented in the fourth column of the figure 49 above. The transliteration is justified for the sake of better

«V'olo neke mulo, nda v'ova mbo nda v'a lupale, tuyevite ča lwa o manu vapopya v'eti, omālā v'e tali ka vakwete e sumbilo. Ka vasumbile o manu la va kwavo čapyālā nģo e nene u ndinga vī vakači lo kulinga po kati ko winģi. Pwāi tupopya e ti ha vo si ko a mālēhē va čituwa momo vana vakači v'e tavo ka valinģi ovi na vi a čo o vyo.»⁴⁰

The proposals above are important to be introduced within the effort of the government to insert native languages into the education system, in that they reveal the gaps left with respect to the orthographical decisions made to write the Umbundu language. Now that Umbundu has to be taught the proposals are coated with enough justification to be implemented in the didactical material and other important documents which may be used as classroom material. At the same time, it is important to consider the expected reactions – in general, negative ones – of those who have memorised the church-initiated orthographical guidelines. Innovation often encounters opposition of many kinds and from very influential agents who may not take a minute to try an understanding of the reasons for such endeavours, mainly when coming from younger “unknown” agents.

This chapter has attempted to present a proposal of the harmonization of the Umbundu orthography. First of all it revisited the most important publications which constituted the major influence on the way Umbundu is written today. Then, the work looks at slightly and confusingly different standard of Umbundu: the PSO, the CSO, the JWSO and the SSO. Next, a proposal of the orthographic harmonization is made at the level of the alphabet by looking at the most confusing letters making it possible to conform to the use of the alphabet to the principles of simplicity or economy, the principle of convenience and the principle of conformity. Finally, a proposal of a more disjunctive approach to the Umbundu orthography is made so as to facilitate teaching and learning

³⁹ The same translation done in the footnote 32 above.

⁴⁰ Below is the English translation of the passage in SSO:

Nowadays, be it in the villages, be it in the cities, we hear people saying that today's children do not show respect. They do not respect people, mainly with regard to the criminal actions that they perpetrate in the communities. However, not all of them show these kinds of behavior because those who go to church do not behave badly.

comprehension by deconstructing some aspects of the previous research and so contribute to the understanding of why a more disjunctive approach to the Umbundu orthography at the morphosyntactic level is to be taken into account.

The following chapter will present relevant conclusions and recommendations of this study by looking at the aspects of the research project's objectives, the constituency of the different chapters and the way the parts of the research work support the arguments of the thesis. Besides, the chapter will advance recommendations for decision makers and other stakeholders to take into account in the labour of the development of the Umbundu language as a means by which communication takes place in the classroom in a coherent way.

6. MAJOR RESEARCH CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCHERS

The present chapter deals with the conclusive part of the research work, taking into account, on the one hand, the description of the objectives of the research, the overall assumptions and limitations underlying the study; the overall contribution to the understanding of the multilingual nature of the southern Africa States, its processes and implications; the language planning research and its contribution to the understanding of language planning in Angola; the contribution of corpus planning research to the understanding of the orthographic harmonization of the Umbundu language and the pertinence of the proposal presented for the harmonization of the Umbundu orthography for educational purposes. On the other hand, recommendations will be presented regarding the target population; the education agents; the critical look at the previous research and its importance for the understanding of the State of the Umbundu orthography standards, and a recommendation is aimed to the openness to new ideas so that the Umbundu orthography is better understood and made apt for educational purposes.

This research work considers the following assumptions:

1. Corpus planning is a complex undertaking that demands regular revision that considers the historical factors, updating and constant adoption from various sources, and adaptation of existing rules and technological facilities in order to make instruction easier.
2. The consideration that the orthography of the Umbundu language has been designed by people who, having not had enough linguistic knowledge of the language, used the resources (knowledge and technology) under their disposal making the orthography look the way it is today. As such, many of the principles and rules underpinning today's Umbundu orthography demonstrate to have been hampered with many misunderstandings and agents (mainly the CSO and the PSO adepts constituting the majority of teachers today) who have the power to study and maintain it do not show any sign of agreement .
3. It is possible and necessary to undertake a research work that can result in a reflection for an all-inclusive discussion and (re)design of an orthography that can help overcome the multiple complications in writing the Umbundu language and, so contribute with a

uniform model for all the users, independent of any religious and institutional prejudices.

4. The respondents of this study can be considered as an acceptable representative sample of the users of the Umbundu language and can be taken as models of native speakers and writers of the Umbundu language.

It is, therefore, crucial to remind of the objectives of this study and explain them in terms of their importance:

- “To revise and discuss the various contexts, factors and motivations that made decision-makers embark on language planning and policy and corpus planning, for the education systems mainly in the southern African context and how they carried out this process at different levels to be successful”. This objective is important in that it pursues the critical study of common characteristics of the Bantu languages which Umbundu shares; verify how much the language of each colonizer has influenced the codification of the bantu languages under their control and the reasons for the codification of the bantu languages, so that a comparison with the Umbundu language can take place. Moreover, this objective also helps see how much the codifications made to other languages in other countries has influenced the making of the Umbundu orthography and in which objective or subjective conditions other languages were thought to be implemented them the education systems.
- “To describe the attempts made, so far, by the Angolan government towards the integration of African languages (Umbundu included) into the education system”. This way, the various processes of native languages’ inclusion and development at various levels, mainly at the education level can be revealed through relevant the assessment of documentation, including legislation and use by various agents since the independence of the country.
- “To provide a historical background of the procedures underpinning the emergency, development and implementation of the Umbundu corpus planning and orthography”. With this objective different agents and contexts contributing to the development of the Umbundu orthography, be they church members at the religious context or independent researchers seeking scientific investigation or politicians trying to capture the necessary support and message comprehension on the part of the inhabitants of the regions where Umbundu is a native language. This also involves the comparative

study of the Umbundu orthography as implemented according to both religion and education background

- “To describe the current State of the Umbundu orthography, by analyzing both existing orthographic systems and identify aspects in which conflicts arise to be overcome, on the basis of proper research tools (questionnaires, written samples from different sources and other constructed synopsis) of data collection”. Here, an attempt is made to scrutinize the various ways the Umbundu orthography is approached, be they Protestant, Catholic or any other and try to reconcile the various forms of writing Umbundu, so that there can be only one standard orthography to follow and use for classes.
- “To collect and register the opinions, feelings, perceptions and attitudes of the adherents of the religious groups about the State of the Umbundu orthography”. This way, users are given voice to express their views and opinions about the Umbundu orthography and why they would adhere to one or another standard orthography. Also, it gives hints to how much Umbundu is used and how important it is for its users.
- “To carry out a critical analysis of the various scholars’ contributions which may have influenced the way the Umbundu orthography has been constructed and propose an improved and harmonized orthography for the current needs of the Angolan education system”. This objective is twofold. On the one hand, a focus is put on the pathways of the Umbundu orthography by tracing the stages of its development with time; see the most preeminent influencers and the legitimacy behind their endeavor. On the other hand, an attempt is courageously made to offer, as a result of imagination and creativity, not only a harmonized, coherent orthography adjusted to be simpler to read and write.

Naturally, this study has experienced various limitations which, in one way or another, have affected the course of it, although not at the point of impeding it to successfully come to an end. On the emotional level, there has always been the solitude that such an enterprise would cause on the researcher. This aspect influences productivity, making a person live like a caged entity overthrown to its fate. Besides, solitude always affects the state of mood by making stress take control of the researcher.

On the financial level, there have been very many hindrances linked to lack of scholarship which would make a huge difference in the expenses. Without sufficient financial support studies

become more limited, sources unachievable and food equilibrium lacks. Consequently, the capacity to think accurately and immunity diminish, affecting production and belates the research work.

As to the bibliographical aspect, the research prior to this one has shown to be very scarce. In fact, there have been very few publications about the orthography of the Umbundu language in Umbundu or in Portuguese. Much of what is available is limited to a few grammars and dictionaries, most of which are either written in foreign languages (mainly in English, French and Germany) and only one publication is totally written in the CSO with very little innovation (Chimbinda, 2015). In terms of language (corpus) planning, the Angolan context has experienced very little since the independence in 1975, only counting on individual initiatives of priests, pastors and relevant church mission members.

The problem with the orthography of Umbundu has elicited two important questions, just to remind of what was stated in the Introduction. The first one is “what factors have caused the emergence and use of different orthographies for Umbundu?” The second question is “how to harmonize these apparently confusing orthographic systems of the same language in the same country with the same speakers for serving education purposes”?

So, it can be concluded that the southern African region where Angola is located shares a common nature of multilingual countries wherein many languages have to cohabit and influence each other. For the majority of these countries the post-independence era did not consider a coherent language policy and planning which would include and guarantee a consequent maintenance of local languages by means of the education system. This is the case of Angola, and Mozambique where the problematic of orthography for non-european languages remains a tabu.

In Angola, the process of language planning has obeyed a political agenda based on a “one people, one nation” moto led by the political party which proclaimed the independence in 1975. Although the Angolan autoctonous languages have been used in the media (public radio broadcast and television) since the beginning of the independence, their inclusion into the education system has still been at the experimental level, demonstrating not enough political interest in the decisive integration of languages like Umbundu in the education system.

Corpus planning as part of language planning and/or language policy has served as a crucial background without which the understanding of this research work would be incomplete in that it encompasses important studies on orthographies of the world, their nature and how different agents at different levels conduct the task of organising and influencing initiatives and decisions on the way

languages are written. The accumulated knowledge in this sub-area has direct implications for this study in terms of:

1. Informed historical registration of the way the Umbundu language orthography was made an object of interest within the overall study of the Bantu languages.
2. Tasks of codification and the standardization of *corpus* carried out by scholars prior to the study of the Umbundu language in other contexts.
3. The role of the missionaries in the decisions made to write the Umbundu language and the purpose for which they decided to encode the Language.
4. The principles underlying orthography design and the way different agents in different contexts, at different levels, embarked upon the choice and use of writing systems.
5. The implications of technology and specialization in orthography for the codification, use and maintenance of the Umbundu language.

The results from the questionnaires testify, to a large extent, that the respondents are divided into those who would prefer the CSO and those who would prefer the PSO. Moreover, there is a tendency for the respondents to prefer the PSO in detrimental of the CSO.

6.1. Main conclusions of the study

The major conclusions of this study, under the evidences produced with regard to the Umbundu orthography, are as follows:

1. Members of both the Catholic and the Protestant churches write Umbundu according to the Standard Orthography (SO) in use in the Bible of their denomination. This indicates that most of the ability to write the Umbundu language is still influenced by the religions of the respondents. In other words, a Catholic respondent writes Umbundu in a way that may, to a large extent, complicate a Protestant respondent. Besides, the study has revealed that many influencers (mainly missionaries) came from different parts of the world with preconceived orthographical principals that they applied to the coding of Umbundu as evidenced in the figures rightly presented in this work.
2. The data show that there is perceived motivation, on the part of the church members, to study Umbundu in the public schools. This information sustains the need to work on the harmonization of the Umbundu orthography to, extensively, make it uniform for all the Umbundu speaking community, independent of being church goers or agnostic citizens.

3. Most of the knowledge of Umbundu on the part of the respondents is acquired either in the home environment or in the church environment. This means that the importance that the church and the family give to Umbundu is not to be despised.
4. There is a tendency of the PSO to be preferred for the education system. This tendency can be explained by the fact that the first translations of the Bible were made by Protestants as shown in the data (see appendix 20 and 22), mainly influenced by previous studies of other languages in other countries of Africa, such as Tanzania, South Africa and so forth, whose missionaries were generally Protestants. Besides, most of the ruling party members who had been educated in Protestant missions. So, they may have had greater influence on which standard Umbundu orthography some State related documents should be written.
5. There is a clear discrepancy both between the members of the two main churches and among the members of the same church with regard to the way they write the same words of Umbundu which shows that there has not been a focused and formal training of the Umbundu language as it should be in any language teaching/learning programme. Each person writes Umbundu as they understand that should be.
6. The CSO and the PSO are different in many ways. The same applies to the differences shown both in the JWSO and SSO. There is a tendency in them to compete in terms of whose way of Umbundu writing is most effective, whose Umbundu orthography is easier to read and write and which one should be adopted for the education system. This involves the ego between those who feel to be the first ones to establish Umbundu writing with the mass – the Protestant ones – and the ones who feel to be the most educated and thus most prepared to determine the intellectual course of the shape of the orthography – the Catholic others. Finally, this has influenced to a greater extent on the variation of Umbundu orthographic standards as seen in the material in the Appendixes 5 to 16 and in the tests in the Appendixes 15 to 27.
7. Teachers/trainers have an acceptable academic background to understand the phenomenon of various orthographic standards happening with the Umbundu language and this may help specialists to create a common platform at the level of corpus planning towards a coherent harmonization of the Umbundu orthography. Additionally, a few initiatives, with potential to give substance to a coherent endeavour into the establishment of a consensual orthography for the education system, have been advanced by some local linguists.
8. There is a great will, on the part of the teachers/trainers, to overcome the current problem of the existing standards of Umbundu in the education system. This demonstrated will may serve

as a creator of synergies to organise symposia, seminars and conventions on a regular basis towards dealing with the harmonization of the Umbundu orthography without any misconceptions. Being in the frontline in the education sector, trainers are valuable influencers to students and the members of the community they serve.

9. Programmes and coursebooks are still scarce which makes most of the respondents draft their own material for students. Material design still depends heavily on export which makes it difficult to reach the hands of its right consumers in the classroom. Besides, there is a clear relaxation on the part of decision-makers in the sector of education as to whether that is a wise idea to introduce the native languages into the education system, an attitude which derives from the idea that in a “one-people-one-nation” State that Angola has been conceived for about 47 years, resorting to native languages would trigger tribalist sentiments in the users.
10. Teachers consider resorting to alphabets other than the Portuguese one important for the amelioration and harmonization of the Umbundu orthography. In fact, initiatives have been engendered to try out letters from other writing systems to accommodate to the needs of the Umbundu languages today. For example, some Czechoslovakian alphabet letters have proven to be economical alternatives to some confusion in the way some sounds should be better coded. Besides, the UNICODE software offers more possibilities to make the orthographical design simpler to learn and use.

The proposal of the harmonization of the Umbundu orthography has been thought to be a critical appreciation of the most important developmental stages that the Umbundu orthography has undergone in the hands of different agents who shaped the way it looks like nowadays. In this respect, the following conclusions have been reached:

1. The Umbundu orthographies today have been a product of crucial 110 years of elaborations by different agents from, at least, 1854 to 1964 (unless otherwise proven), period of which the most important decisions of what standards of the Umbundu language should be used were made. This has made it difficult for users to renounce prior learnt standards and orthographic practice of theirs in detriment of more functional and harmonized ones. In general, people tend to resist new trends no matter how functional they may appear.
2. Many of the agents who contributed to the design of the Umbundu orthographies are missionaries who were not necessarily specialists in language studies, including priests, pastors, and individual researchers who understood the importance of knowing and noting down the aspects of this important language to the best of the intellectual and technical

capacities they have had. Again, native speakers have not been taken into account to help detect the imprecisions resulting from non native limitations.

3. The Umbundu language orthography is made of multiple standards, the principal of which are the PSO, the CSO, the JWSO and the SSO currently in simultaneous use within the regions where the Ovimbundu people inhabit. Apart from these standards there are other ways that influential people – mainly politicians and religious leaders –tend to use and force other to recognize and follow. This amalgamation of writing approaches has demonstrated to lead to a difficult way to cope with Umbundu in written form in various institutions. This has, for example, resulted in wrongly written toponymical and anthroponymical systems, causing serious embarrassments and/or damage which have long been ignored by decision-makers whose native knowledge is somehow scarce.
4. The morphosyntactic aspect of the Umbundu orthography has, on the one hand, been impaired by what is here believed to be a probable misunderstanding of what is called “the classes of nominal prefixes” by the immeasurable and influential work of the German linguist Wilhelm Heinrich Immanuel Bleek (8 March 1827 – 17 August 1875) whose systematization makes it difficult to detach the irregular forms of the verb “to be” in Umbundu (“oku kala”) from the nouns they collocated with, making these various forms of its conjugations be inaccurately taken, it is here believed, as nominal prefixes. This has resulted in a complicated and contradictory presentation of a list of 18 noun prefix class whose use has proven much forced and dependent on regular memorization of rules. On the other hand, the adoption of a more conjunctive approach to the Umbundu orthography has made it even worse for its potential to make different grammatical classes be written as one complex-word.

6.2. Suggestions with regard to the population, educational agents and previous research

The suggestions of this study encompass the target population, the educational agents, the critical look at the previous research in orthography and its implications on the harmonization of the Umbundu writing system.

Regarding the target population, it can be suggested that a consciousness action plan has to be decided by the educational, political and religious agents together in a democratic way, in order to negotiate, explain and agree on the best ways to make the orthography of Umbundu harmonized for all the users independent of the religions they belong to. This suggestion will only be feasible if the

Government and the Umbundu users are all involved and their points of view, perceptions and opinions are taken into due consideration and that the scientific stance is coherently and convincingly passed over to the consumers. Moreover, it is important to secure that the teaching of Umbundu starts as early as possible and encouraged to be widely used in print in public places and institutions in the regions where it is used.

As to the educational agents, it is suggested that the most important educational institutions such as ISCED/Benguela, Jean Piaget Institute and the higher teacher training schools in Benguela and other provinces, carry out research, coordinate and engage in cooperative work with regard to the Umbundu language programme design. Moreover, this undertaking should consider the involvement of an interdisciplinary commission made of historians, anthropologists, sociolinguists, psycholinguists, grassroots, politicians, the clergy and the like, as well as consider regularly organized workshops and seminars in which members of language related departments should be summoned up to contribute.

Another important aspect of this suggestion is the guaranteeing of institutional support towards the concretization of the project of an orthographic accord which will serve as a framework through which documentation of various kinds shall be elaborated. This will imply the existence of a State body responsible for the coordination of the whole programme on a regional basis, the involvement of IT (Information Technology) specialists to respond to the technical aspects of programming and operationalization of special cases of the digital area, such as making the keyboard respond to newly suggested letters.

Having had a critical look at the results from the questionnaires and the documents analysis, it can be proposed that,

1. Agents involved in the use and teaching of Umbundu should be encouraged to study and propose practical changes found necessary to ameliorate the Umbundu orthography so as to facilitate the processes of teaching and learning. One of these practical changes can be informed by the research proposals of this study and others which are related to innovative initiatives.
2. Seminars and workshops should be organized to discuss and decide where, in the orthographic arrangements of Umbundu, to adopt a disjunctive or a conjunctive approach.
3. A deeper study of the "System of noun prefixes" for the Bantu languages based on Bleek (1869) who refers to the forms of the verb "to be" (oku kala) as prefixes of the nouns, adjectives and prepositions to which they are attached, should be firmly carried out. This

recommendation helps decide whether to approach these morphosyntactic arrangements in a conjunctive or in a disjunctive way.

6.3. Suggestions for future work

This study cautions that future researchers of the Umbundu orthography should take consciousness of the existing possibilities offered by the technological advances in terms never developed before. So, openness to new trends may help both researchers and Umbundu users resort to less complicated ways of writing the Umbundu language for the benefit of the users, mainly the students of this language and help on law, signposts and pamphlets' translations and interpretations for the community. A cooperative approach to the Umbundu orthography design – which can make Umbundu language teachers, trainers and users meet with computer technicians in regular skills-sharing encounters – should be promoted, so that computer software at the keyboarding level is developed to make the writing of Umbundu easier.

Moreover, a classroom material design taskforce should be created with the purpose of sketching the elaboration of textbooks based on the recommendations of this research work, considering an experiential period to be decided towards the implementation of more functional ways of writing the Umbundu language that is simpler, more economical and reader/learner-friendly. This will help participants understand the various ways the Umbundu orthography appears on paper, as demonstrated in the appendices' section related to tests and exams below.

As it can be seen, this research cannot be taken as an exhaustive substitute of the previous ones, but it can be taken as a serious reflection on the matter of Umbundu orthography. As such, the nearest step forward encompasses a campaign to promote, in the shortest run, funds to hold the first regional convention of the Umbundu orthography with the sole purpose of conglomerating, consulting and sketching a potentially agreed standard orthography of the Umbundu language, one that can break the existing limiting sectarian and religion-driven paradigm.

REFERENCES

- Ager, D. E. (2005). Prestige and Image Planning. In E. Hinkel (Ed.). *A Handbook of Research in Second Language Teaching and Learning*. (pp. 1035 – 1054). Laurence Erlbaum Associates, Inc.
- Algeo, J. (2010). *The Origins and developments of the English language (6th edition)*. Wadsworth, Cengage Learning.
- Alpatov, V. M. (2017). Scripts and politics in the USSR. *Studi Slavistici xiv*, 9-19. [https://doi: 10.13128/Studi_Slavis-21936](https://doi.org/10.13128/Studi_Slavis-21936)
- Alves, A. (1951). *Dicionário etimológico Bundo-Português (Ilustrado com muitos milhares de exemplos, entre os quais 2.000 provérbios indígenas), 2*. Tipografia Silvas, Lda. & Centro Tipográfico Colonial.
- Andrade, C.A. (1962). Two expressions of angolanity. *Presence Africaine*, 42/43, 14/15, 70–84. <http://casacomum.org>
- Andrzejewski, B. W. (1983). Language reform in Somalia and the modernization of Somali vocabulary. In Fodor, I. & Hagège, C. (Eds.). *Language Reform: History and future*, 69 – 84. Buske.
- Angola. Resolução n°3 de 23 de maio de 1987. (1987). <https://www.axl.cefan.ulaval.ca/afrique/Angola-loi-1987-3.htm>.
- Angola. (2001). Lei n.º 13/01 de 31 de Dezembro de 2001. I série - n.º 65. *Diário da República*.
- António, A., Cuianda, L.N. & Bonifácio, R. (2012). *Manual de Alfabetização Umbundu: Elivulu lyokulilongisa okutanga kwenda okusonēha Umbundu*. África International.
- Asher, R. E. & Simpson, J. M. Y. (1994). *The Encyclopedia of Language and Linguistics*. Pergamon.
- Assembleia Nacional/Imprensa Nacional. [https://www.unicef.org › unicef.org.angola](https://www.unicef.org/unicef.org.angola).
- Ayto, J. (1983). English: Failures of language reforms. In I. Fodor, & C. Hagège (Eds.). *Language reform: History and future*. (pp. 85 – 100). Buske.
- Babane, M.T. & Chauke, M. T. (2015) Cross-border Language Harmonization: A Case of Xitsonga (South Africa) and Shangani (Zimbabwe). *Journal of Sociology and Social Anthropology*, 6(3): (pp. 435 – 440). Kamla-Raj.
- Baker, T.L. (1994). *Doing social research (2nd Ed.)*. McGraw-Hill Inc.
- Baldauf, R. B. (2004). Language Policy and Planning: Recent trends, future directions. Paper presented at the AAAL, Portland, OR, May 2004. <http://espace.library.uq.edu.au/view/UQ:24518>

- Baldauf, R. B. (2012). Introduction – Language Planning: Where have we been? Where might we be going? *Revista Brasileira Linguística Aplicada*, 12(2), 233-248.
- Barnitz, J.G. (1978). Interrelationship of orthography and phonological structure in learning to read (Technical Report No. 57). Urbana Champaign, IL: University of Illinois at Urbana-Champaign. *Center for the Study of Reading*. (ERIC Document Reproduction Service No. ED 150 546). <https://www.ideals.illinois.edu/bitstream/handle>
- Baroni, A. (2011). Alphabetic vs. non-alphabetic writing: Linguistic fit and natural tendencies. *Rivista di Linguistica*, 23(2), 127-159.
- Batibo, H. (2000). The sounds of Africa: their phonetic characteristics. In V. Web & Kembo-Sure (Eds.). *African voices: an introduction to the languages and linguistics of Africa*. (pp. 133 -159). Oxford University Press.
- Batsikama, P.C.M. (2015). *Nação, nacionalidade e nacionalismo em Angola* [Unpublished doctoral thesis] Universidade Fernando Pessoa. <http://bdigital.ufp.pt/bitstream/10284/5191/1/Tese%20Patr%C3%ADcio.pdf>
- Bendor-Samuel, J. & Hartell, R. L. (eds.) (1989). *The Niger-Congo languages*. University Press of America.
- Bijeikienė, V. & Tamošiūnaitė, A. (2013). *Quantitative and qualitative research methods in Sociolinguistics*. Vytautas Magnus University Press.
- Bitchener, J. (2010). *Writing an applied linguistics thesis or dissertation: a guide to presenting empirical research*. Macmillan.
- Bleek, H. I. (1869). *Comparative Grammar of the South African Languages*. Trübner.
- Bontrager, G. H. (2015). *Concepts and issues in orthographic design*. [Unpublished master's thesis]. University of Florida. Available at <https://www.researchgate.net/publication/283015919>. DOI: 10.13140/RG.2.1.5001.0326
- Botswana. (2011). *Population and housing census 2011 dissemination seminar (2014)*. Gaborone. Statistics Botswana.
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27 – 40. <https://www.unicef.org/unicef.org.angola>. DOI 10.3316/QRJ0902027.
- Bussmann, H. (2006). *Routledge Dictionary of Language and Linguistics*. Taylor & Francis e-Library.
- Cabré, M. T. (1999). *Terminology: Theories, methods and applications*. (Edited by Juan C. Sager & translated by Janet Ann DeCesaris). John Benjamins.

- Cacumba, J.S.C. (2020). *Competências de Literacia Académica Bilingue (Português/Inglês) dos Graduandos em Ensino da Língua Inglesa em Angola*. [Unpublished Doctoral Thesis]. University of Minho.
- Campbell, G. L. & King, G. (2011). *The Concise Compendium of the World's Languages. (2nd Edition)*. Routledge.
- Capello, H. & Ivens, R. (1881). *De Benguella as terras de lacca, vol. 2*. Imprensa Nacional.
- Central Intelligence Agency. (2009). Namibia. In *The World Factbook*. <https://www.cia.gov/library/publications/the-world-factbook/>
- Childs, G. T. (2003). *An Introduction to African Languages*. John Benjamins B. V.
- Chimbinda, J. S. F. (2015). *Okuluka Kwetu*. ETU.
- Chimbutane, F. & Benson, C. (2012). Expanded spaces for Mozambican languages in primary education: Where bottom-up meets top-down. In McCarty, T. & Hornberger, N. (Eds.). *Globalization from the Bottom up: Indigenous Language Planning and Policy in Globalizing Spaces and Places. International Multilingual Research Journal*, 6(1), 8–21. <https://www.tandfonline.com/doi/abs/10.1080/19313152.2012.639278?journalCode=hmrj> 20
- Clark, J. B.; Haque, E.; & Lamoureux, S. A. (2012). The role of language in processes of internationalization: Considering linguistic heterogeneity and voices from within and out in two diverse contexts in Ontario. *Theorizing International Education: (Shifting) contexts, concepts, methods*: 41, (3), (5). <https://doi.org/10.5206/cie-eci.v41i3.9212>
- Cluver, A. D. V. (1992). Language planning models for a post-apartheid South Africa. *Language problems and language planning*, 16(2), pp. 105-136.
- Cobbarrubies, J., & Fishman, A. (1983). *Progress in Language Planning: International Perspectives*. De Gruyter.
- Constituição da República de Angola, de 5 de Fevereiro de 2010. (2010). I série – N° 23, *Diário da República*. <http://tribunalsupremo.ao/uploads/2018/05/c...>
- Cooke, G. (2021). Orthography. In J. Stanlaw (ed.). *The International Encyclopedia of Linguistic Anthropology*. John Wiley & Sons, Inc. <https://onlinelibrary.wiley.com>
- Cooper, R. (1989). *Language planning and social change*. Cambridge University Press.
- Cornelis, S. (1991). Stanley au Service de Léopold II: La Fondation de l'Etat Indépendant du Congo (1878-1885) ". In S. Cornelis (Ed.), H.M. Stanley: *Explorateur Au Service du Roi*. (41–60). Royal Museum for Central Africa.

- Coulmas, F. (2005). *Sociolinguistics: The study of speaker's choices*. Cambridge University Press.
- Coulmas, F. (Ed.). (1998). *The handbook of sociolinguistics*. Blackwell Publishing.
- Coupland, N. & Kristiansen, T. (2011). SLICE: Critical perspectives on language (de)standardization. In T. Kristiansen & N. Coupland (Eds.). *Standard Languages and Language Standards in a Changing Europe*. (11-35). Novus.
- Craven, M. (2015). "Between law and history: the Berlin Conference of 1884-1885 and the logic of free trade". *London Review of International Law*, 3, 31–59. [https://www.google.com/search?q=Craven%2C+M.+\(2015\)...](https://www.google.com/search?q=Craven%2C+M.+(2015)...)
- Crystal, D. (2003). *The Cambridge encyclopedia of language (2nd edition)*. Cambridge University Press.
- Da Costa, T. M. C. J. (2015). *Umbundismos no Português de Angola: Proposta de um Dicionário de Umbundismos*. [Unpublished doctoral thesis]. FCSH, Universidade Nova de Lisboa.
- Dalby, A. (2004). *Dictionary of languages: The definitive reference to more than 400 languages*. A & C Black Publishers Ltd.
- Daniel, H. E. (2002). *Ondisionaliu Yumbundu/Dicionário de Umbundo: Umbundo – Português*. Edições Naho.
- Daniels, P. T. (2001). Writing Systems. In A. Mark & J. Rees-Miller (Eds.). *The Handbook of Linguistics*, (pp. 43 – 80). Blackwell.
- Daoust, D. (1997). Language Planning and Language Reform. In F. Coulmas (Ed). *The Handbook of Sociolinguistics*. (pp. 436 – 452.). Blackwell Publishers.
- Darquennes, J. & Vandenbussche, W. (2015). The standardization of minority languages – introductory remarks. *Sociolinguistica*, 29(1), 1-16. <https://doi.org/10.1515/soci-2015-0003>
- Darquennes, J. (2013). Language policy and planning in indigenous language minority settings. *Revue Française de Linguistique Appliquée*, 18 (2), pp. 103 – 119.
- Dash, R. K. (2019). Ecolinguistics and endangered languages. *IOSR Journal of Humanities and Social Science (IOSR-JHSS)*, 24(5), pp. 14-16.
- Davies, W. V. & Ziegler, E. (2015). *Language planning and Microlinguistics: From policy to interaction and vice versa*. Palgrave Macmillan.
- De Blij, H.J. & Muller, P. O. (1997). *Geography: Realms, Regions, and Concepts*. John Wiley & Sons, Inc.
- De Oliveira, A. J. (2012). *O Grande Livro dos Provérbios Angolanos*. União dos Escritores Angolanos.
- De Oliveira, S. (2015). A Implantação do Modelo Formal de Ensino em Angola (Séculos XV-XX). *Revista Brasileira de História e Educação*, 15(2), p. 55-80.

- Decreto Lei n° 62 de 6 de Abril de 1978 sobre a Criação do Instituto de Línguas Nacionais (ILN), do Ministério da Educação. (1978). <https://www.lexlink.eu/legislacao/angola/14793/ia-serie/por-tipo-de-documentolegal/>
- Decreto n° 40 de 18 de Novembro d 1985 sobre a Criação do instituto nacional de línguas nacionais. (1985). I Série, n° 92. <https://www.lexlink.eu/legislacao/angola/14793/ia-serie/por-tipo-de-documentolegal/>
- Decreto Lei n.º 17/16 Lei de Bases do Sistema de educação e Ensino de 7 de Outubro de 2016. (2016). I série – N° 170. <https://www.unicef.org> › [unicef.org](https://www.unicef.org).angola › files
- DeFrancis, J. (2006). The prospects for Chinese writing reform. *Sino-Platonic Papers*, 171. www.sino-platonic.org.
- Deprez, K.; Du Plessis, T. & Teck, L. (2001). *Multilingualism, the Judiciary and Security Services: Belgium, Europe, South Africa, Southern Africa*. Van Schaik.
- Deumert, A. & Vandebussche, W. (2003b): Standard languages: Taxonomies and histories. In A. Deumert & W. Vandebussche, (Eds.). *Germanic standardizations. Past to present*. (1-14). John Benjamins.
- Diarra B. (2003). Choice and description of national languages with regard to their utility in literacy and education in Angola. In: A. Ouane (Ed.). *Towards a Multilingual Culture of Education*. (333–348). *UNESCO Institute for Education*.
- Do Nascimento J. P. (1894). *Grammatica do Umbundu ou Lingua de Benguella*. Imprensa Nacional.
- Doiz, A.; Lasagabaster, D. & Sierra, J. (2011). Internationalization, multilingualism and English-medium instruction. *World Englishes*. 30(3), pp. 345-359. <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-971X.2011.01718.x>
- Dovalil, V. (2020). Process of destandardization and demotization in the micro – macro perspective: The case of Germanic languages. In G. C. Kimura & L. A. Fairbrother (Eds.). *A language management approach to language problems: Integrating macro and micro dimensions*. (177 – 198). John Benjamins Publishing Company. <http://languagemanagement.ff.cuni.cz/node/352>
- Dwyer, A. M. (2005). *The Xinjiang conflict: Uyghur identity, language policy, and political discourse*. East-West Center Washington.
- Eastman, C.M. (2001). National language/Official language. In R. Mesthrie (Ed.), *Concise Encyclopedia of sociolinguistics* (pp. 657–662). Pergamon.
- Eberhard, D. M., Simons, G. F. & Fennig, C. D. (eds.). (2019). *Ethnologue: Languages of the World*. (22nd digital edition). SIL International. <http://www.ethnologue.com>

- Elbro, C. (2005). Literacy acquisition in Danish: A deep orthography in cross-linguistic light. In R. M. Joshi & P. G. Aaron (Eds.). *Handbook of orthography and literacy*. <https://www.routledgehandbooks.com/doi/10.4324/9780203824719.ch3>.
- Elkartea, G. (2010) *Language standardization: Basque recovery II*. Creative Commons.
- Ennis, M. W. (1962). *Umbundu: Folk Tales from Angola*. Beacon Press/Soares.
- Ferguson, C.A. & Heath, S.B. (1981). Glossary. In C.A. Ferguson & S.B. Heath (Eds.). *Language in the USA* (pp. 527–533). Cambridge University Press.
- Fettes, M. (1997). Language planning and education. In R. Wodak & D. Corson (Eds.). *Language policy and political issues in education*. (pp. 13 – 22). Kluwer Academic.
- Firmino, G. (2005). *A questão linguística na África pós-colonial. O caso do Português e das línguas autóctones em Moçambique*. Texto Editores.
- Fishman, J. A. (2006). *Do not leave your language alone: The hidden status agendas within language planning in language policy*. Lawrence Erlbaum.
- Frost, R. (2005). Orthographic systems and skilled word recognition processes in reading. In M. J. Snowling & C. Hulme (Eds.). *The science of reading: A handbook*. (pp. 272 – 295). Blackwell.
- Fry, P. (2001). Introdução. In P. Fry. (Org.). *Moçambique: Ensaios* (pp. 13 – 20). Editora da UFRJ.
- Gaio, G. (2016). Angolanidade e angolanização: Interfaces da governança no pós-guerra civil em Angola. *Odeere: Revista do Programa de Pós-Graduação em Relações Étnicas e Contemporaneidade – UESB, 1(1)*, 266 – 290. Gall, M.D.
- Ghyselen, A., Delarue, S. & Lybaert, C. (2016). Studying standard language dynamics in Europe: Advances, issues & perspectives. *Taal en Tongval. 68(2)*, 75-91. https://www.researchgate.net/publication/308318930_Studying_standard_language_dynamics_in_Europe_Advances_issues_perspectives
- Gonzalez, A. (2004). Language planning and intellectualisation. *Current issues in language planning. 3(1)*, 5 – 27.
- Goswami, U. (2010). A psycholinguistic grain size view of reading development across languages. In N. Brunswick, S. McDougall & P. de Mornay Davies (Eds.). *Reading and dyslexia in different orthographies*. (pp. 23 – 42). Psychology Press.
- Goundar, P. R. (2017). The characteristics of language policy and planning research: An overview. *Sociolinguistics - interdisciplinary perspectives*. <https://www.intechopen.com/books/sociolinguistics-interdisciplinary-perspectives/the-characteristics-of-language-policy-and-planning-research-an-overview>.

- Gries, S. T. (2009). What is Corpus Linguistics? *Language and Linguistics Compass*, 3: 1
https://www.researchgate.net/publication/220531593_What_is_Corpus_Linguistics
- Guthrie, M. (1948). *The classification of the Bantu languages*. International African Institute.
- Hambly W. D. (1934). The Ovimbundu of Angola. *The Anthropological Series, Vol. XXI, Vol. 329(2)*, 93–356.
- Hartshorne, K. (1987). Language policy in African education in South Africa 1910 – 1985. In D. Young (Ed.). *Bridging the gap between theory and practice in English second language teaching: essays on honour of I. W. Lanham*. (pp. 62 – 80). Maskew Miller Longman.
- Haugen, E. (1966b) *Language conflict and language planning: the case of Modern Norwegian*. Harvard University Press
- Haugen, E. (1972). *The ecology of language*. Stanford University Press.
- Haugen, E. (1987). *Blessings of Babel*. Mouton de Gruyter.
- Haugen, E.(1966a). Dialect, Language and Nation. *American Anthropologist* , 68(4), 922-935.
- Heale, R. and Forbes, D. (2013). Understanding triangulation in research. *Evid Based Nurs*, 16(14), 1 - 13. <https://www.isg.pt/wp-content/uploads/2021/03/2072-Article-Text-2133-1-10-20210226-SCOPUS.pdf>
- Heine, B. (1992). Language policies in Africa. In R. Herbert (Ed.). *Language and society in Africa: the theory and practice of sociolinguistics*. Witwatersrand University Press.
- Hermans, A. (1991). Language planning and terminology planning. Sociolinguistic aspects. *Meta*, 36 (4), 687–689. <https://doi.org/10.7202/003068ar>.
- Hérnard, F.; Diamond, L. and Roseveare, D. (2012). Approaches to internationalization and their implications for strategic management and institutional practice. *OECD/IMHE Higher Education Programme*. 11(12).
<https://www.oecd.org/education/imhe/Approaches%20to%20internationalization%20-%20final%20-%20web.pdf>
- Heugh, K. (1987b). Trends in language medium policy for a postapartheid South Africa. In D. Young (ed.). *Language: planning and medium in education, (206 – 220)*. The Language Education Unit and SAALA.
- Hill, L. (2010). Language and status: On the limits of language planning. *Stellenbosch Papers in Linguistics*, 39(1), pp.41-58
https://www.researchgate.net/publication/274558849_Language_and_status_On_the_limits_of_language_planning

- Holmes, J. (2008). *Introduction to sociolinguistics (3rd edition)*. Longman.
- Holmes, J. & Hazen, K. (Eds.). (2014). *Research methods in Sociolinguistics: A practical guide*. Wiley Blackwell.
- Hornberger, N. H. (2006). Frameworks and models in language policy and planning research. In T. Ricento (Ed.). *An introduction to language policy: theory and method*. (pp. 24 - 41). Blackwell.
- Hornberger, N.H., & King K. A. (1999). Authenticity and unification in Quechua language planning. In S. May (ed). *Indigenous community-based education* (pp. 160 – 80). *Multilingual Matters*.
- Hosken, M. (2003). Creating an orthography description. *orthog_desc.sxw*. https://scripts.sil.org/cms/scripts/render_download.php?format=file&media_id=OrthographyDescPaper&filename=OrthographyDescPaper.pdf
- Hungwe, K. (2007). Language policy in Zimbabwean education: historical antecedents and contemporary issues. *Compare*. 37(2): 135– 149.
- IECA (2009). *Hinário Evangélico da Igreja Evangélica Congregacional em Angola* (6ª edição). Centro de Publicações Cristãs.
- INE. (2016). *Censo 2014: Recenseamento geral da população e habitação (Resultados definitivos)*. INE.
- Inverno, L. (2011). *The restructuring of Portuguese morphosyntax in interior Angola - evidence from Dundo (Lunda Norte)*. [Unpublished Doctoral Thesis]. Universidade de Coimbra.
- Jaffe, A. & Walton, S. (2000). The voices people read: Orthography and the representation of non-standard speech. *Journal of Sociolinguistics*, 4(4), pp. 561 – 587.
- Jimbi, B. I. & Sicala, D. V. (2020). Empreendedorismo ortográfico: uma proposta de harmonização e preservação da língua Umbundu falada no centro-sul de Angola. *Revista Digital de Políticas Lingüísticas*, 13(13), 115 – 138. <https://revistas.unc.edu.ar/index.php/RDPL/article/view/31560/32289>
- Jimbi, B. I. (2018). A reflection on the Umbundu corpus planning for the Angola education system: towards the harmonization of the Catholic and the Protestant orthographies. *Actas do XIII Congreso Internacional de Lingüística Xeral, Vigo 2018*, 475-482.
- Jones, M. C. (Ed.). (2015). *Policy and planning for endangered languages*. Cambridge University Press.
- Jorge, M. (1998). *Para compreender Angola: Da política à economia*. Dom Quixote.
- Kajibanga, V. (2000). *A alma sociológica na ensaística de Mário Pinto de Andrade: Uma introdução ao estudo de vida e obra do primeiro sociólogo angolano*. I.N.I.C.

- Kamwangamalu, N. M. (2011). Language planning: approaches and methods. In E. Hinkel (Ed.). *A handbook of research in second language teaching and learning. (Volume II)*. (pp. 888 – 904). Routledge.
- Kaplan, R. B. (2011). Macro language planning. In E. Hinkel (Ed.). *A handbook of research in second language teaching and learning. (Volume II)*. (pp. 924 – 935). Routledge.
- Kaplan, R. B. & Baldauf, Jr. R. B. (2012). *Language and language-in-education planning in the pacific basin*. Kluwer Academic Publishers.
- Kaplan, R. B. & Baldauf, R. B. (1997). *Language planning from practice to theory*. Clevedon, Philadelphia, Toronto Sydney. Multilingual Matters Ltd.
- Karan, E. (2006). Writing system development and reform: a process. [Unpublished Master Dissertation]. University of North Dakota.
- Katzenellenbogen, S. (1996). It didn't happen at Berlin: Politics, economics and ignorance in the setting of Africa's colonial boundaries. In P. Nugent & A. I. Asiwaju (Eds.). *African Boundaries: Barriers, Conduits and Opportunities*, (pp. 21 – 34). Pinter.
- Kheng, C. C. S. & Baldauf Jr, R. B. (2011a). Micro language planning. In E. Hinkel (Ed.). *A handbook of research in second language teaching and learning. (Volume II)*. (pp. 936 – 951). Routledge.
- Kheng, C. C. S. & Baldauf Jr, R. B. (2011b). Global language: (de)colonisation in the new era. In E. Hinkel (Ed.). *A handbook of research in second language teaching and learning. (Volume II)*. (pp. 952 – 969). Routledge.
- Kioko, A. N., Njoroge, M. C., & Peter Mburu Kuria P. M. (2012). Harmonizing, the orthography of Gĩkũyũ and Kĩkamba, in N. O. Ogechi, J. A. N. Oduor, & P. Iribemwangi (eds.). *The harmonization and standardization of Kenyan languages: Orthography and other aspects*, Cape Town. (pp. 39 – 63). CASAS.
- Kloss, H. (1969). *Research possibilities on group bilingualism: A report*. International Center for Research on Bilingualism.
- Koelle, S.W. (1854). *Polyglotta Africana: Or, a comparative vocabulary of nearly three hundred words and phrases in more than one hundred distinct African languages*. Church Missionary House.
- Küster, M. W. (2019). Open and closed writing systems. Some reflections. In Y. Haralambous (Ed.). *Graphemics in the 21st Century. Brest, June 13–15, 2018. Proceedings (Grapholinguistics and its Applications, Vol. 1)*. (pp. 17 – 26). Fluxus Editions.

- Lanham, L. (1978). An outline history of the languages of Southern Africa. In L. Lanham & K. Prinsloo (Eds.) *Language and communication studies in South Africa*. (pp. 13 – 28). Oxford University Press.
- Le Guennec, G. and Valente, J. F.(1972). *Dicionário Portugues-Umbundu*. Instituto de Investigação Científica de Angola.
- Lecomte, E. (1899). *Ondaka ia suku ou Doutrina Christa em umbundu e portuguez*. Imprensa Nacional.
- Legère, K., Trewby, R., & van Graan, M. (2000). *The implementation of the Namibian language policy in education: Lower primary grades and pre-service teacher education*. Namibia Education Research Association.
- Leimgruber, J. (2019). *Language planning and policy in Quebec: A comparative perspective with views from Wales and Singapore*. Narr Francke Attempto.
- Lewis, M. P. (ed.), (2009). *Ethnologue: Languages of the World, (6th digital edition)*. SIL International. <http://www.ethnologue.com/16>
- Liddicoat, A. J. (2005). Corpus planning: Syllabus and materials development. In E. Hinkel (Ed.). *A handbook of research in second language teaching and learning*. (pp. 993 – 1012). New Laurence Erlbaum Associates, Inc.
- Linn, A., Sanden, G. R., & Piekkari, R. (2018). Language standardization in sociolinguistics and international business: Theory and practice across the table. In T. Sherman & J. Nekvapil (Eds.). *English in business and commerce: Interactions and policies; English in Europe*. 5(14), (pp. 19 – 45). Aalborg: Mouton de Gruyter. <https://vbn.aau.dk/en/publications/language-standardization-in-sociolinguistics-and-international-bu>.
- Lo Bianco, (2010). Language policy and planning. In N. H. Hornberger & S. L. McKay (Eds.). *Sociolinguistics and language education*, (pp. 146 – 176). Multilingual Matters.
- Lodhi, A. (1993). The language situation in Africa today. *Nordic Journal of African Studies*, 2(1), 79–86. <http://www.njas.helsinki.fi/pdf-files/vol2num1/lodhi.pdf>.
- Longtau, S. R. (2014). Practical perspectives on minimum essentials for effective orthography designs in Nigeria. *Jos Linguistics Circle Meeting, Mountain View, NTA Jos*, 27 de Fevereiro de 2014. https://www.academia.edu/6917694/Practical_minimum_essentials_for_design_of_effective_orthographies_in_Nigeria.
- Luansi, L. (2003). Angola: movimentos migratórios e estados pré-coloniais: Identidade nacional e autonomia regional. *International Simposium Angola on Move, Transport Routes*,

- Communications and History*. Berlin: Lukonde Luansi, 23–24.
[https://www.zmo.de/angola/Papers/Luansi_\(29-03-04\).pdf](https://www.zmo.de/angola/Papers/Luansi_(29-03-04).pdf)
- Lüpke, F. (2011). Orthography development. In P. K. Austin & J. Sallabank (Eds.), *The Cambridge handbook of endangered languages*, (pp. 312–336). Cambridge University Press.
- Madiba, M. (1999). Multilingualism and nation-building in the “new” South Africa: Problems and issues. *Languages matter: studies in the languages of Africa*, 30(1), 59 – 82.
- Maho, J. F. (2009). NUGL Online: The online version of the New Updated Guthrie List, a referential classification of the Bantu languages, 4 June, 2009.
<http://goto.glocalnet.net/mahopapers/nuglonline.pdf>.
- Malone, S. (2004). *Manual for developing literacy and adult education programs in minority language communities*. UNESCO. Available at:
<http://unesdoc.unesco.org/images/0013/001351/135164e.pdf>
- Mangoya, E. (2009). Divergent approaches to corpus processing: The need for standardization. *Lexikos 19 supplement (AFRILEX-reeks)*, 19(1), 26-37.
- Mateus, M.H.M. & Villalva, A. (2007). *Linguística*. Editorial Caminho.
- Matiki, A. J. (2002). Language planning and Linguistic exclusion in the Legislative Process in Malawi. *World Congress on Language Policies*. [Conference paper]. http://www.linguapax.org/wp-content/uploads/2015/07/CMPL2002_T1_AMatiki.p
- May, S. (2005). Language policy and minority language rights. In E. Hinkel (Ed.). *A handbook of research in second language teaching and learning*. (pp. 1055 – 1074). Laurence Erlbaum Associates, Inc.
- McArthur, T. (Ed.).(1992). *The oxford companion to the English language*. Oxford University Press.
- McEney, T. and Hardie, A. (2011). *Corpus linguistics: Methods, theory and practice*. Cambridge University Press.
- McEney, T. and Wilson, A. (2001). *Corpus Linguistics: An introduction (2nd Edition)*. Edinburgh University Press.
- McEntee-Atalianis, L. J. (2016). A network model of language policy and planning: The United Nations as a case study. *Language problems & language planning*. 40(2), 187-217.
<https://www.semanticscholar.org/paper/A-network-model-of-language-policy-and-planning%3A-as-McEntee-Atalianis/cb4894fc1dce974756b574ad77ed8bbda2152716>
- Mezzadri, M. (2003). *L'italiano essenziale: testo di grammatica per studenti stranieri dal livello principianti (A) al livello intermedio-alto (B2)*. Guerra edizioni.

- Miethaner, U. (2000). Orthographic transcriptions of non-standard varieties: The case of earlier African-American English. *Journal of sociolinguistics*, 4(4), 534 – 560.
- Miller, P., Kargin, B. & Guldenoglu, T. K. B. (2012). Differences in the reading of shallow and deep orthography: developmental evidence from Hebrew and Turkish readers. *Journal of Research in Reading*, 0(0), 1–24. <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9817.2012.01540.x>
- Mkhabela, B. R., Nxumalo, Z. G., & Bhebhe, S. (2018). Using an Indigenous Language as a Medium of Instruction in Primary Schools in the Kingdom of Eswatini: Insights and Views from Parents. *International Journal of Informatics, Technology & Computers*, 1(1), 126 – 138.
- Mokibelo, E. (2014). The national language as a language of instruction in Botswana primary schools. *Language and Education*, 28 (5), 21-435. <https://minerva-access.unimelb.edu.au/bitstreams/fe3dd219-c85f-5b86-966e-3c95a8a5a7f1/download>
- MONTEIRO, Domingas, Henriques. Tradições Nacionais e Identidades: Recolha e Estudo de canções Festivas e de óbito Kongo e Ovimbundu, 2014. Disponível em: <<https://repositorio-aberto.up.pt/bitstream/10216/74367/2/32353.pdf>>.
- Mordaunt, O. G. (1990). Swaziland's Language Policy for Schools. *Educational Studies*, 16 (2): 131 – 140
- Moser, G. (1983). Héli Chatelain: Pioneer of a National Language and Literature for Angola. *Research in African Literatures, Special Issue on African Literary History*, 14(4), 516 – 537.
- Mühlhäusler, P (2000). Language planning and language ecology. *Current issues in language planning*, 1 (3), 306 – 367.
- Murphy, M. L. (1996). Dictionaries and Orthography in Modern Africa. *Lexikos*, 6 (6), 44 – 70.
- Mwakikagile, G. (2010). *Nyerere and Africa: End of Era*. New Africa Press.
- Nahir, M. (1984). Language planning goals: A classification. *Language problems and language planning*, 8(3), 294–327.
- Ndimande-Hlongwa, N. (2010). Corpus planning, with specific reference to the standard isiZulu in the media. *Alternation: Interdisciplinary journal for the study of the arts and humanities in southern Africa*, 17(1), 207-224.
- Nekvapil, J., and Sherman, T. (2015). An introduction: Language management theory in language policy and planning. *International journal of the sociology of language*, 232, 1–12. 10.1515/ijsl-2014-0039. <https://www.degruyter.com/document/doi/10.1515/ijsl-2014-0039/html>

- Neto, J. P. (2011). Os Governos de Norton de Matos em Angola. *Boletim Da Sociedade De Geografia De Lisboa, Série 129*- N.os 1-12 Janeiro – Dezembro, 87 – 106.
- Neto, M. (2012). *In Town and Out of Town: A social history of Huambo (Angola) 1902-1961*. [Unpublished doctoral thesis]. University of London.
- Ngcobo, M. (2007). Language planning, policy and implementation in South Africa. *Glossa: an Ambilingual interdisciplinary Journal*. 2(1): pp. 1 – 23
- Ngcobo, M. (2007). Language planning, policy and implementation in South Africa. *Glossa: An ambilingual Interdisciplinary Journal*, 1(2), 156 – 169.
- Nguaiko, N. E. (2015). *The New Otjiherero Dictionary English Herero Otjiherero Otjiingirisa*. Namibia: Authorhouse.
<https://dicg5fhpqu3ab.cloudfront.net/gkzjbmsnhulvywfrpxeqditoaac8c9bd3b8e6.pdf>
- Noré, A. and Adão, A. (2003). O Ensino Colonial aos “Indigenas” de Angola. Antecedentes do Ensino Rudimentar Instituído pelo Estado Novo. *Revista Lusófona de Educação*, 001. 101 – 126.
- Nurse, D. and Tucker, I. (2001). A Survey Report for the Bantu Languages. SIL International.
<https://www.sil.org/system/files/reapdata/82/45/16/8245165576250264113481067886>
- Odendaal, G. (2013a). Restandardization defined as democratising language planning. *Stellenbosch papers in linguistics plus*, 42(1), 183-203,
<https://spilplus.journals.ac.za/pub/article/view/165>
- Oko, C. M. W. (2018). Orthography development for Darma (the case that wasn't). *Language documentation and conservation*, 12(1), 15–46.
- Osborn, D. (2010). *African languages in a digital age: Challenges and opportunities for indigenous language computing*. HSRC Press.
- Ottenheimer, H. J. (2012). Ideology and orthography: dictionary construction and spelling choice in the Comoro islands. *Études océan Indien [online]*, 48, <http://journals.openedition.org/oceanindien/1521>; DOI: 10.4000/oceanindien.1521.
- Ouane A. (Ed.) (2003). *Towards a Multilingual Culture of Education*. UNESCO Institute for Education.
- Palomé, E. C. (1967). *Swahili language handbook*. Center for Applied Linguistics.
- Pedro J. D. et al. (2013). *Harmonização Ortográfica das Línguas Bantu de Angola (Kikongo, Kimbundu, Umbundu, Cokwe, Mbunda, oshikwanyama)*. The Centre for Advanced Studies of African Society & Instituto de Línguas Nacionais do Ministério da Cultura.
- Phillipson, R. (1992). *Linguistic imperialism*. Oxford University Press.

- Pillièrre, L.; Andrieu, W.; Kerflec, V. and Lewis, D. (2018). Norms and margins of English. In L. Pillière, W. Andrieu, V. Kerflec & D. Lewis (Eds.). *Standardising English: Norms and margins in the history of the English language*. (pp. 3 – 21). Cambridge University Press.
- Pinto, A. O. (2014). Representações Culturais da Rainha Njinga Mbandi (c.1582-1663) no Discurso Colonial e no Discurso Nacionalista Angolano. *Estudos Imagética*. Rio de Janeiro: UERJ / CH-FLUL.
<http://repositorio.ul.pt/bitstream/10451/12027/1/Representa%C3%A7%C3%B5es%20Culturais%20da%20Rainha%20Njinga%20Mbandi%20com%20imagens.pdf>
- Pinto, H. & Silva, A. A. (2021). Língua umbundu: caminhos para a sua preservação. *RAC: Revista Angolana de Ciências*. 4(1). <https://doi.org/10.54580/R0401.05>
- Pitman, B. (1905). *A plea for alphabetic reform*. Krehbiel Co.
- Polit, D.F., Beck, C.T. and Hungler, B.P. (2001). *Essentials of Nursing Research: Methods, Appraisal and Utilization (5th Ed.)*. Lippincott Williams & Wilkins.
- Ponso L. C. (2011). Políticas linguísticas actuais em Angola e Moçambique: O modelo monolíngua do estado nacional europeu X o modelo plurilíngua das nações Africanas. *The XI Luso Afro Brazilian Congress of Social Sciences, August 2011*. Bahia Federal University, pp. 7–10.
- Protopapas, A. & Vlahou, E. L. (2009). A comparative quantitative analysis of Greek orthographic transparency. *Behavior research methods*, 41(4), (991-1008)
<https://link.springer.com/article/10.3758/BRM.41.4.991>
- Queiroz, A. (1978). *Angola – A via agreste da liberdade*. Artur Queiroz & Ulmeiro para a Língua Portuguesa.
- Quinta, J., Brás, J. & Gonçalves, M. N.(2017). O Umbundo no poliedro linguístico angolano: a Língua Portuguesa no entrelaçamento do colonialismo e póscolonialismo. *Revista Lusófona de Educação*, (35) 1, 137-154.
- Rasinger, S. M. (2008). *Quantitative research in Linguistics: An introduction*. Continuum.
- Reagan, T. G. (2010). *Language Policy and Planning for Sign Languages*. Gallaudet University Press.
- Ricento, T. (2000). Historical and theoretical perspectives in language policy and planning. *Journal of Sociolinguistics* 4(2), 196-213.
- Ricento, T. (2006) Language policy: Theory and practice – An introduction. In T. Ricento, (Ed.). *An introduction to language policy: Theory and method*. (pp. 3 – 9). Blackwell Publishing.
- Ricento, T. & Hornberger, N. H. (1996) Unpeeling the onion: Language planning and policy and the ELT professional. *TESOL quarterly*, 30(3), 401-28.

- Richards, J. C. & Schmidt, R. (2010). *Longman dictionary of language teaching and applied linguistics (4th digital edition)*. Pearson Education Limited.
- Roberts, D. (2014). Review of “developing orthographies for unwritten languages”. *Language documentation & conservation*, 8(1), 781–787 <http://hdl.handle.net/10125/24625>.
- Roberts, D., Basnight-Brown, D. & Vydrin, V. (2019). Marking tone with punctuation: orthography experimentation and reform in eastern Dan (Côte D'Ivoire). In Y. Haralambous (Ed.). *Graphemics in the 21st Century. Brest, June 13–15, 2018. Proceedings (Grapholinguistics and Its Applications, Vol. 1)*. (pp. 293–327). Fluxus Editions.
- Rocha, D. (2011, May 6). Images Of Nzinga Mbandi Ngola's Diplomacy, in Luanda, in 1621: History, Embossings, and Narrative (Pepetela). *III Encontro Nacional de Estudos da Imagem 03 a 06 de maio de 2011 - Londrina - PR, 847 - 857*. <http://www.uel.br/eventos/eneimagem/anais2011/trabalhos/pdf/Denise%20Rocha-20UNITINS,%20Palmas..pdf>.
- Rogers, M. (2015). *Specialized translation: Shedding the 'non-literary' tag*. Palgrave Macmillan.
- Rosenberg, S. Weisfelder, R. F.; Frisbie-Fulton, M. (2005). *Historical dictionary of Lesotho*. Scarecrow Press.
- Ruiz, R. (1995). Language planning considerations in indigenous communities. *The Bilingual Research Journal*. Winter, 19(1), 71–81. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.125.8096&rep=rep1&type=pdf>.
- Ryan, F., Coughlan, M. and Cronin, P. (2009). Interviewing in qualitative research: The one-to-one interview. *International Journal of Therapy and Rehabilitation*, 16(6), 309 – 314.
- Sacalembe, J. E. (2021). O silenciamento da língua angolana Umbundu. https://www.google.com/search?q=J%C3%BAlio+Epalanga+Sacalembe2&rlz=1C1GCEA_enUS843US843&oq=J%C3%BAlio+Epalanga+Sacalembe2&aqs=chrome..69i57.6480033j0j7&sourceid=chrome&ie=UTF-8
- Sanders W. H. and W. E. Fay (1885). *Vocabulary of the Umbundu Language, Comprising Umbundu-English and English-Umbundu*. Beacon Press.
- Santos, B. S. (2018). *O fim do império cognitivo*. Coimbra: Almedina.
- Sasoma, J. M. (2015). Interferências da língua umbundu na língua portuguesa na região de Benguela: estudo descritivo. [Unpublished Master's thesis]. Universidade Nova de Lisboa. <http://hdl.handle.net/10362/15797>.

- Schadeberg, T. C. (1982) Nasalization in Umbundu. *Journal of African Languages and Linguistics* 4, 109-132.
<https://scholarlypublications.universiteitleiden.nl/access/item%3A2727531/download>
- Schadeberg, T. C. (1986). Tone cases in Umbundu. *Africana Linguistica*, 10, 423 – 447.
<https://scholarlypublications.universiteitleiden.nl/handle/1887/8816>
- Schadeberg, T. C. (1990). *A Sketch of Umbundu*. Rüdiger Köppe Verlag.
- Schiffman, H. F. (1996). *Linguistic culture and language policy*. Routledge.
- Schleef, E. and Meyerhoff, M. (2010). “Sociolinguistic methods for data collection and interpretation”.
 In M. Meyerhoff, & E. Schleef, (Eds). *The Routledge Sociolinguistics Reader*. (pp1–17).
 Routledge.
- Schroeder, L. (2008). *The Bantu orthography manual: for linguistic and literacy fieldworkers*. SIL International.
- Sgall, P. (1987). Towards a theory of phonemic orthography. In P. A. Luelsdorff, (Ed.). *Orthography and phonology*. (pp. 1 – 30). John Benjamins Publishing Company.
- Siiner, M., Hult, F. M. & Kupisch, T. (2018). Situating language acquisition planning. In M. Siiner, F. M. Hult, & T. Kupisch, (Eds.). *Language policy and language acquisition planning* (P. 1). Springer.
- Silva, A. M. (2008). *Norton de Matos e a Auto-defesa face à História Coeva e Futura*. Pombalina/Coimbra University Press.
- Simões, A., Sacanene, B., Iriarte, A., Almeida, J.J. & Macedo, J. (2020) Towards a Morphological Analyzer for the Umbundu Language. *SLATE*.
<https://drops.dagstuhl.de/opus/volltexte/2020/13023/pdf/OASlcs-SLATE-2020-10.pdf>
- Simons, G. (1994). Principles of multidialectal orthography design. *Notes on Literacy*, 20(2), 13 – 34.
- Simons, G. F. & Fennig, C. D. (eds.). (2018). *Ethnologue: Languages of Africa and the World. (21st edition)*. SIL International. <http://www.ethnologue.com>.
- Soares, M. C. & Agostinho, M. B. (2016). A Coleção Ovimbundu do Museu Nacional, Angola 1929-1935. *MANA*. 22(2): 493 – 518, <http://dx.doi.org/10.1590/1678-49442016v22n2p493>.
- Spolsky, B. (2012). What is language policy? In B. Spolsky, (Ed.). *The Cambridge Handbook of Language Policy*. (pp. 3 – 15). Cambridge University Press.
- Steigertahl, H. (2018). Similarities and differences between varieties of speech: Expressions of language ideologies and attitudes in four Namibian communities. *SOAS Working Papers in Linguistics*. Vol. 19 (2), (pp. 133-153). SOAS University of London.

- Stewart, W. A. (1968). A sociolinguistic typology for describing national multilingualism. In J. A. Fishman (Ed.). *Readings in the sociology of language*. (pp. 531 – 545). Mouton Publishers, The Hague.
- Storch, A. (2013). Language and ideology. In Lüpke, F. and Storch, A. (Eds.). *Repertoire and choices in African languages*. (pp. 123 - 175). The Gruyter Mouton.
- Stover, W. M. (1887). *Otuikanda tuokufetika lokutanga Umbundu*. West Central African Mission.
- SUZI (2013). *Zungumza Kiswahili: Speak Swahili*. Educational Books Publishers Ltd.
- Tagliamonte, S. A. (2006). *Analysing sociolinguistic variation*. Cambridge University Press [Kindle Edition].
- Taylor-Leech, K. (2016). From practice to theory: The contributions of Richard B. Baldauf Jr. to the field of language policy and planning. *Current Issues in Language Planning*, 17(1), 1 – 10. <https://www.tandfonline.com/doi/full/10.1080/14664208.2016.1129477>
- Teixeira, J. (2014). *Como funcionam as línguas. Uma iniciação às ciências da linguagem*. Edições Húmus.
- Tonkin, A. (2015). Language planning and planned languages: How can planned languages inform language planning? *Interdisciplinary Description of Complex Systems*, 13(2), 193 – 199, https://indecs.eu/2015/indecs_13_2.pdf
- Torrend J. (1891). *A Comparative Grammar of the South-African Bantu Languages*. Kegan Paul, Trench, Trübner & Co., Ltd.
- UNDESA. (2017). World Population Prospects: The 2017 revision, key findings and advance tables. *working paper No. ESA/P/WP/248*. United Nations.
- UNESCO. (1996, 6th – 9th June). World conference on linguistic rights: Barcelona Declaration, Executive Board, 150th[Conference paper]. <https://unesdoc.unesco.org/ark:/48223/pf0000104267>.
- Unger, J. M. (1996). *Literacy and script reform in occupation Japan*. Oxford University Press.
- UNISA. (2011). *MSLPLPV Tutorial Letter 101*. Unisa.
- Valente J. F. (1964). *Gramática Umbundu: A língua do centro de Angola*. Instituto de Investigação Científica de Angola.
- Van den Berghe, P.L. (1968). Language and 'nationalism' in South Africa. In J. A. Fishman, C. A. Ferguson & J. das Gupta (Eds.). *Language problems of developing nations*, (pp. 215 – 224). John Wiley and Sons.

- Van Dyken, J. R. and Lojenga, C. K. (1993). Word boundaries: Key factors in orthography development. In R. L. Hartell (ed.). *Alphabets of Africa*, (pp. 3–20). UNESCO & SIL.
- Van Els, T. (2005). Status planning for learning and teaching. In E. Hinkel (Ed.). *A handbook of research in second language teaching and learning*. (pp. 971 – 992). Laurence Erlbaum Associates, Inc.
- Van Teijlingen, E. R. and Hundley, V. (2002). The importance of pilot studies. *Nursing Standard: Official Newspaper of the Royal College of Nursing*, June 2002, https://www.researchgate.net/publication/11173521_The_Importance_of_Pilot_Studies
- Venezky, R. L. (2004). In search of the perfect orthography. *Written language & literacy*, 7(2), 39–63, <https://www.semanticscholar.org/paper/In-search-of-the-perfect-orthography-Venezky/ec3b825feb8fd24c8c0aa8d79929bdb3d40038c1>
- Vosters, R. and Villa, L. (Eds.) (2015). *The historical sociolinguistics of spelling (special issue of written language & literacy, 18(2))*. John Benjamins.
- Webster, N. (1967, 1789). *Dissertations on the English language*. The Scholar Press Ltd.
- Wedgwood, M. A. H. (1856, 2009). *A dictionary of English etymology, Vol. I. (A – D)*. Trübner & CO. /Brittle Books Project.
- Werner, A. (1919). *Introductory Sketch of the Bantu Language*. Kegan Paul, Trench, Trubner & Co., Ltd.
- Westermann, D. & Ward, I. C. (1957). *Practical Phonetics for Students of African Languages*. Oxford University Press, International African Institute.
- WFP. (2019). Angola interim country strategic plan (2020–2022). https://executiveboard.wfp.org/document_download/WFP-0000108616
- Wiley, T. G. (1996.). Language planning and policy. In S. L. McKay & N. H. Hornberger (Eds.). *Sociolinguistics and language teaching*. (pp. 103 – 148). Cambridge University Press.
- Willems, H. (2014). *Historical and archeological aspects of Egyptian funerary culture: Religious ideas and ritual practice in Middle Kingdom elite cemeteries*. BRILL. https://books.google.com/books/about/Historical_and_Archaeological_Aspects_of.html?id=zwYJBAAAQBAJ
- Wright, S. (2012). Language policy, the nation and nationalism. In B. Spolsky (Ed.) *The Cambridge handbook of language policy*. (pp. 59 – 78). Cambridge University Press.
- Yifan, W. (2019). What are we calling “latin script”? Name and reality in the grammatological terminology. In Y. Haralambous(Ed.). *Graphemics in the 21st Century. Brest, June 13–15,*

2018. *Proceedings (Grapholinguistics and Its Applications, Vol. 1)*. (pp 91 – 109). Fluxus Editions.
- Zarnikhi, A. (2016). *Towards a systemic model for terminology planning*. Cambridge Scholars Publishing.
- Zau, D.G.N. (2011). A língua portuguesa em Angola: Um contributo para o estudo da sua nacionalização. [Unpublished doctoral thesis] Beira Interior University. <http://hdl.handle.net/10400.6/1844>.
- Zau, F. (2005). O professor do ensino primário e o desenvolvimento dos recursos humanos em Angola (Uma visão prospectiva). [Unpublished doctoral thesis]. Open University. <https://repositorioaberto.uab.pt/handle/10400.2/2468>.
- Zhao, S. (2011). Actors in language planning. In E. Hinkel (Ed.) *A handbook of research in second language teaching and learning (Volume II)*, (pp. 905 – 923). Routledge.

APPENDICES

Appendix 1: Questionnaire on the harmonization of the Umbundu orthography

QUESTIONNAIRE ON THE HARMONIZATION OF THE UMBUNDU ORTHOGRAPHY QUESTIONÁRIO SOBRE A HARMONIZAÇÃO DA ORTOGRAFIA DO UMBUNDU In English and Portuguese/em Inglês e Português

Dear respondent,

Today, the Angolan government has been focusing attention on the insertion of Angola's native languages, usually called 'national languages', into the education system. Umbundu is an Angolan native language widely known to be spoken in Benguela, Huambo, Huila, and Bié provinces. The problem is that two written versions of the Umbundu language have been currently put into force: One is the Catholic version, and the other is the Protestant (Evangelical) one. By completing this questionnaire, we will be able to understand the current State of the Umbundu orthography and the need to harmonize the two versions, towards the facilitation of teaching and learning. This research will be conducted and reported to obtain the Ph.D. degree in Sociolinguistics at the University of Minho. We would be thankful if you could answer the following questions honestly and we assure that the information from you will be taken confidentially.

N.B.: By answering this questionnaire you are freely committing yourself to this useful research by allowing us to work with the data that you provide. To guarantee anonymity, you do not have to sign in any way.

Yours sincerely

Botelho Isalino Jimbi

Caro inquerido,

É sabido que, actualmente, o governo Angolano tem direccionado atenções na inserção das línguas nativas Angolanas, conhecidas como "línguas nacionais". O Umbundo é uma língua falada principalmente nas províncias de Benguela, Huambo, Huila, e Bié. O problema é que duas versões escritas do Umbundo têm sido usadas: uma versão Católica e a outra Protestante. Ao completares este questionário, ajudar-nos-ás a compreender o estado da ortografia do Umbundu e a necessidade actual de harmonização das duas versões, de modos a facilitar o processo de ensino e aprendizagem. Esta pesquisa será levada a cabo e apresentada para a obtenção do grau de Doutoramento em Sociolinguística na Universidade do Minho. Agradecemos que respondesses as perguntas abaixo com honestidade, e asseguramos que a informação prestada será tratada muito confidencialmente.

N.B.: Ao responderes este questionário estarás a marcar, de livre vontade, um útil compromisso com esta pesquisa, autorizando-nos a trabalhar com a informação que nos fornecer. Para garantir-se o necessário anonimato, não tens que deixar quaisquer assinaturas, de modo algum.

Respeitosamente,

Botelho Isalino Jimbi

IMPORTANT INSTRUCTION/(INSTRUÇÃO IMPORTANTE):

- A. Circle the lines you choose out of the proposed ones. (Assinale com círculo as alíneas sugeridas que escolheres.)
- B. Use the line spaces to justify your choices as convenient. (Justifique convenientemente nos espaços deixados, caso escolhas opções que o exijam.)

GROUP ONE(GRUPO UM): PERSONAL BACKGROUND (INFORMAÇÃO PESSOAL)

1. What do you do? (Qual a tua profissão?)
- a. Pastor. (Pastor)
 - b. Priest. (Padre)
 - c. Catholic member
 - d. Protestant member
 - e. Other. Please, specify. (Outra. Especifique) _____.
2. Indicate your age according to the categories below. (Indique a sua idade nas categorias abaixo.)
- a. 10 to 18 (10 a 18)
 - b. 18 to 25 (18 a 25)
 - c. 25 to 30 (25 a 30)
 - d. 30 to 40 (30 a 40)
 - e. Over 40 (Mais de 40)
3. Which province have you been living for the time being? (Em que província vives neste momento?)
- a. Benguela
 - b. Huila
 - c. Huambo
 - d. Bié
4. Where were you born?(Onde é que nasceste?) _____
5. Indicate the language(s) you can speak. (Indique a (s) língua (s) que falas)
- a. Portuguese (Português)
 - b. Umbundo
 - c. Ambas (Both)
6. Write the religious confession of your own or of your family. (Qual a tua religião ou da família?)
- a. Protestant (Evangélica)
 - b. Catholic (Católica)
 - c. Other. Please, specify. (Outra. Especifique) _____.
7. Write the name(s) of language(s) your mother speaks more at home.
(Mencione os as línguas que a tua mãe fala mais em casa)
- a. Portuguese (Português)
 - b. Umbundu.
 - c. Other. Please, specify. _____
(Outra. Por favor, Especifique)
8. Indicate whether your mother can write in the language(s) that follow.
(Indique se a tua mãe sabe escrever na(s) língua(s) que se seguem)
- a. Portuguese.
 - b. Umbundu.
 - c. Other. Please, specify. _____.

(Outra. Por favor, Especifique.)

9. Write the name(s) of language(s) your father can speak.
(Escreva o(s) nome(s) da(s) língua(s) que o teu pai fala.)

10. Indicate whether your father can write in the language(s) that follow.

(Indique se o teu pai aprendeu a escrever na(s) língua(s) que se seguem.)

- a. Portuguese (Português)
- b. Umbundu
- c. Both (Ambas)

Other. Please, specify. (Outra. Por favor, Especifique.) _____.

11. If you have a family, what language do you speak at home?

(Em caso de viveres em família, que língua falas em casa?)

- a. Portuguese (Português)
- b. Umbundu

c. Other. Please, specify. (Outra. Por favor, Especifique.) _____.

GROUP TWO: LEARNING THE LANGUAGE. (GRUPO DOIS: APRENDER A LÍNGUA)

12. Have you ever studied an African language of Angola other than Portuguese at public school?

(Já alguma vez estudaste uma língua Africana de Angola que não seja o português na escola pública?)

- a. Yes. (sim)
- b. No. (Não)

13. If yes, can you write and read in that language? (Se sim, sabes escrever e ler nessa língua?)

- c. Yes. (Sim)
- d. No. (Não)

14. Studying the Umbundu language is (Estudar a língua Umbundu é:)

- a. A waste of time nowadays. (Perda de tempo hoje em dia)
- A way of maintenance (Forma de preservá-la)

- b. A patriotic duty
(Um dever patriótico)

- c. A linguistic right in itself
(Um direito linguístico em si mesmo)

- d. All the above but line(s) (Todos acima menos a(s) alínea(s) _____).

- e. Other. Please, specify. (Outra. Por favor, Especifique.) _____.

GROUP THREE: PRACTICE OF UMBUNDU. (GRUPO TRÊS: UMBUNDU NA PRÁTICA)

15. Umbundu is written in two different versions: the Catholic orthography and the Evangelical one.

(O Umbundu está escrito em duas versões: A versão Católica e a versão Evangélica.)

- a. True (verdade)
- b. False (falso)
- c. Not sure (em dúvida)

16. Indicate the language(s) used in the church where you profess.

(Que língua (s) usada(s) nos cultos que professa?)

- a. Umbundu
- b. Portuguese (Português)

- c. Both (Ambas)
17. Indicate the language(s) you prefer reading the most in the church.
(Em que língua(s) mais preferes ler na igreja?)
- Umbundu
 - Portuguese (Português)
 - Both (Ambas)
18. Do you know of a State school where Umbundu is taught?
(Sabes de uma escola pública em que o Umbundo é ensinado?)
- Yes (Sim)
 - No (Não)
19. Which Standard Umbundu Orthography is used at State school, do you believe?
(Que padrão do Umbundu que acreditas estar em usado na escola estatal?)
- The Catholic (O Católico)
 - The Protestant (O Protestant)
 - Both (Ambos)
 - Other. Please, mention (Outro, por favor menciona)_____
20. Which Standard Umbundu Orthography should be used at State school, do you believe?
(Que Padrão Ortográfico do Umbundu achas que deveria ser usado na escola?)
- The Catholic (O Católico)
 - The Protestant/Evangelic) (O /Protestant/Evangélico)
 - Both (Ambos)
 - None (Nenhum)
 - Other. Please, mention (Outro, por favor menciona) _____
21. Write the reason for your choice in 18. (Escreva a razão da tua escolha em 18?)
- _____
- _____
22. Would you like to have learnt Umbundu at school?
(Gostavas ter aprendido o Umbundo na escola?)
- Yes (Sim)
 - No (Não)
23. Please, comment on the choice made above?
(Por favor, comente a sua resposta acima)
- _____
- _____
- _____
24. Try the most to write the following sentences in Umbundu, so it can be inferred how skillful you are in it:
(Tenta o teu máximo de escrever as seguintes frases na língua Umbundu para que se saiba o quanto podes):
- The baby knows it and knows the Lord. He is a witch. (O bebé sabe e conhece o Senhor. Ele é feiticeiro.)
- _____
- _____
- _____

- b. The King has sprinkled the bag; the queen has sprinkled the outdoors. (O Rei borrifou o saco, a Rainha borrifou no vão.)

- c. The hippopotamus uses to defecate here. (O hipopótamo costuma defecar aqui.)

- d. The snail slithers. (O caracol escorrega.)

- e. Dad has left and I stayed yawning. (O pai partiu e eu fiquei a bocejar. Estou com raiva.)

- f. Misery causes hunchback. (A miséria causa corcunda.)

Thank you very much for your cooperation

Muito obrigado pela tua disponibilidade

Braga, 2020

Appendix 2: Teacher/trainer questionnaire on the Umbundu orthography harmonization

**TEACHER/TRAINER QUESTIONNAIRE ON THE UMBUNDU ORTHOGRAPHY
HARMONIZATION
QUESTIONÁRIO DO FORMADOR/PROFESSOR SOBRE A HARMONIZAÇÃO DA
ORTOGRAFIA DO UMBUNDU**

Dear teacher/trainer,

Umbundu Language Teaching is at its seemingly uncertain experimental stage in the education system and it is believed that teachers have recently been involved in training to respond to this new challenge in Huambo, Huila and Benguela provinces. The problem is that teachers use different orthographies according to whether they belong to either Catholic or Protestant (Evangelic) religion. By filling in this questionnaire we will come to know a) how aware and concerned you are of the difference between the Catholic and the Protestant orthographies and b) what you think should be done to harmonize the two versions for a common orthographic accord for the education system. The resulting data will help us understand the factors behind the existence of the two orthographies and propose a revised and harmonized orthography for the users in general and in the education system in particular. This research will be conducted and reported to obtain the Ph. D. degree in Sociolinguistics with the University of Minho, Portugal. We guarantee that the information from you will be taken in high confidentiality.

N.B.: By answering this questionnaire you are freely committing yourself to this useful research by allowing us to work with the data that you provide. To guarantee anonymity, you do not have to sign in any way.

Yours sincerely

Botelho Isalino Jimbi

Caro(a) professor(a)/formador(a),

O Ensino da Língua Umbundu está na sua quase hesitante fase experimental no sistema educacional. Nas províncias do Huambo, da Huila e de Benguela professores e professoras têm sido preparados (as) a responder à esta nova demanda. O problema que se impõe é a existência de dois padrões ortográficos principais que poderão ser usados pelos professores, dependendo de se eles são membros da Igreja Católica ou se da Protestante. Ao preencheres este questionário ajudar-nos-ás a compreender a) o quanto sabes e te preocupas com a diferença existente entre a ortografia católica e a evangélica e, b) o que achas que deve ainda ser feito para a harmonização das duas versões por uma ortografia de consenso para o sistema de educação. Os resultados desta pesquisa serão relevantes na medida em que nos ajudarão a compreender os factores que sustentaram a existência das duas ortografias e propor uma revisão e uma harmonização ortográficas para os usuários, em geral, e para o sistema de educação, em particular. A pesquisa é feita dentro de um programa de Doutoramento em Sociolinguística na Universidade do Minho, Portugal. Esperamos que respondas as perguntas que se seguem conscientemente, sendo certo que asseguramos que toda a informação será tratada com a máxima discrição.

N.B.: Ao responderes este questionário estarás a marcar, de livre vontade, um útil compromisso com esta pesquisa, autorizando-nos a trabalhar com a informação que nos fornecer. Para garantir-se o necessário anonimato, não tens que deixar quaisquer assinaturas, de modo algum.

Respeitosamente

Botelho Isalino Jimbi

IMPORTANT INSTRUCTION/(INSTRUÇÃO IMPORTANTE):

- C. Circle the lines you choose out of the proposed ones. (Assinale com círculo as alíneas sugeridas que escolheres.)
D. Use the line spaces to justify your choices as convenient. (Justifique convenientemente nos espaços deixados, caso escolhas opções que o exijam.)

1. What is your current professional position? (Please, circle one option). (Qual o teu estatuto profissional. Por favor, circule a alínea a que te enquadras)

- a. Teacher (Professor)
b. Course coordinator (Coordenador de curso)
c. Teacher trainer (Formador de professores)
d. Other. Please, mention. (Outro. Por favor, mencione) _____

2. What is your current academic situation? (Please circle one option)

(Qual a tua situação académica actual? (Por favour, circule uma opção)

- a. *Licenciatura* degree (Licenciatura)
b. MA./MSc. (Mestre)
c. Ph. D. (Doutor)
d. Other. Please, mention (Outra. Por favor, mencione) _____

3. What specialization(s) have you done so far? (Please circle one option)

(Qual/quais a/as tua/tuas especialização/especializações actualmente (por favor, circula uma opção)

- a. Portuguese Linguistics (Linguística/Português)
b. Bantu Linguistics (Linguística Bantu)
c. English Linguistics (Linguística/Inglês)
d. French Linguistics (Linguística/Francês)
e. Other. Please, mention (Outra(s) (Por favor, mencione) _____

4. What is your religion? (please circle one option)

(Qual a tua religião? (Por favour, circula uma opção)

- a. Catholic (Católica)
b. Protestant (Protestante)
c. Other. Please, mention (Other. Por favor, mencione) _____

5. Where were you born? (Please, circle one option)

(Onde é que nasceste? (Por favor, circule uma opção)

- a. In Huambo province
b. In Huila province
c. In Benguela province
d. In other location. Please, mention it. (Outro lugar. Por favor mencione) _____

6. Choose a category that corresponds to your age. (Please circle one option)

(Escolha a categoria da sua idade. Por favor, circule uma opção)

- a. 20 to 30
b. 30 to 40
c. 40 to 50
d. 50 up

7. **Did you learn Umbundu when you were at the primary school? (Please, circle one option). (Estudaste Umbundu na escola primária? Por favor, confirme uma opção)**
- Yes (sim)
 - No (não)
8. **If yes, how well did you learn it? (Please, tick one option) (Se sim, quão bem o aprendeste? Por favor, confirme uma opção)**
- I can write and read it very well (aprendi a escrevê-lo e lê-lo muito bem)
- I can understand and speak it very well (Aprendi a entendê-lo e falá-lo muito bem)
- All the above (Todas as opções acima)
- None of the above (Nenhuma das opções acima)
- I have no full command of it .(Não tenho um domínio total)
9. **If not, how did you come to know it to the point of teaching the future teachers? (Please circle one). (Se não, como é que chegou de conhecê-lo até poder ensiná-lo? Por favor, circule uma opção)**
- It was used at home (Usava-se em casa)
 - It was the language of my community as I grew up.(Foi a língua da minha comunidade)
 - I learnt it at church. (Aprendi-o na Igreja)
 - I have had classes of Umbundu. (Fiz um curso de Umbundu)
 - Others. Please, specify (Outros. Por favour, especifique _____)
10. **Where do you teach Umbundu? (Please, circle one option) (Onde é que ensinas o Umbundu? Por favor, circule uma opção)**
- At ISCED (Higher Institute of Education Science)/Benguela.(No ISCED/Benguela)
 - At Jean Piaget Higher Polytechnic Institute (no Instituto Superior Politécnico “Jean Piaget”)
 - At both.(Em ambos)
 - Other. Please, mention where else (Outro. Por favour, mencione onde mais)
-
11. **How long have you been teaching Umbundu? (Please, circle one option). (Há quanto tempo ensinas o Umbundu? (Por favor, circule uma opção)**
- One year (Um ano)
 - Two years (Dois anos)
 - Three years (Três anos)
 - Five years (Cinco anos)
 - More than five years (acima de cinco anos)
12. **What Umbundu standard orthography do you use? (Please, circle one option). (Que padrão de ortografia da língua Umbundu é que usas? (Por favor, circule uma opção).**
- The Catholic Standard Orthography (O Padrão de Ortografia Católica)
 - The Protestant Standard Orthography (O Padrão de Ortografia Protestante)

- c. Both (Ambos)
- d. None of the above (Nenhum dos acima mencionados)
- e. Other. Please, mention (Outro. Por favor, menciona)

13. **Who designs the curriculum of the Umbundu language teaching? (Please, circle one option).**

(Quem desenvolve o currículo para o ensino do Umbundo? (Por favor, circule uma opção)

- a. Myself (Eu mesmo/mesma)
- b. The State (O estado)
- c. Jean Piaget Institute (O Instituto "Jean Piaget")
- d. Other. Please, mention it. (Outro. Por favor, indique)

14. **Which coursebook(s)/didactical material do you use for classes? (Please, circle one option).**

(Qual o manual/material didático que utiliza para as aulas?) (Por favor, circule uma opção)

- a. Supplied by the government with the title... (Fornecido pelo governo, intitulado...)
Title
(Título) _____
- b. My own compilations (minhas compilações)
- c. None (Nenhum)
- d. Other. Please, mention it. (Outro. Por favor, mencione) _____

15. **Why do you think Umbundu should be taught? (Please, comment succinctly).**

(Porque achas que o Umbundu deve ser ensinado? (Por favor, comente sucintamente).

16. **If you were asked to contribute to the harmonization of the Umbundu orthography for a common *corpus* planning for the education system, what of the aspects bellow would you like to help work out? (Please, circle one option).**

(Se tivesses que contribuir para a harmonização da ortografia do Umbundu por uma planificação de *corpus* para o sistema de educação, que aspectos abaixo gostarias de te dedicar? (Por favor, circule)

- a. Maintain the Catholic writing style.
(Manter a versão ortográfica Católica)
- b. Maintain the Protestant writing style.
(Manter a versão ortográfica Evangélica)
- c. Resort to other languages' alphabets to harmonize the two where necessary).
(Recorrer a alfabetos de outras línguas para harmonizar as duas onde for necessário).
- d. Simplify the sound-to-letter correspondence.
(Simplificar a correspondência letra-som)

e. Suggest an orthographical convention with umbundu-speaking Angolan specialists and researchers. (Sugerir uma convenção sobre a ortografia, com especialistas angolanos/angolanas falantes e estudiosos do Umbundu.)

f. Others. Please, mention which. _____
(Outros. Por favor, indique qual). _____

17. Comment on the reasons for your choice of the options in 15.

(Comente os motivos da sua escolha em 15)

18. Comment on problem(s), if any, that you think there is with Umbundu language teaching and learning?

(Comente a volta de quaisquer problemas, se os houver, que achas que o ensino e aprendizagem do Umbundu enfrentam?)

**THANK YOU SO MUCH FOR YOUR CO-OPERATION
MUITO OBRIGADO PELA COOPERAÇÃO
Braga, 2020**

Appendix 3: Written samples from the teacher/trainer questionnaire on the Umbundu orthography harmonization

Sample 1: Comments on the choice of standard orthography

17 15. Comment on the reasons for your choice of the options in 14.
(Comente a volta dos motivos da sua escolha em 14)

Abreviaturas e siglas das duas variantes para valorizar a cultura e o empenho das que se dedicaram na missão de indagar o que era oral.

THANK YOU SO MUCH FOR YOUR COOPERATION
MUITO OBRIGADO PELA COOPERAÇÃO
Brasá, 2020

17 15. Comment on the reasons for your choice of the options in 14.
(Comente a volta dos motivos da sua escolha em 14)

Abreviaturas e siglas das duas variantes para valorizar a cultura e o empenho das que se dedicaram na missão de indagar o que era oral.

17 15. Comment on the reasons for your choice of the options in 14.
(Comente a volta dos motivos da sua escolha em 14)

Abreviaturas e siglas das duas variantes para valorizar a cultura e o empenho das que se dedicaram na missão de indagar o que era oral.

THANK YOU SO MUCH FOR YOUR COOPERATION
MUITO OBRIGADO PELA COOPERAÇÃO
Brasá, 2020

17 15. Comment on the reasons for your choice of the options in 14.
(Comente a volta dos motivos da sua escolha em 14)

Atendendo a proximidade de que cada letra deve corresponder a um som, arredou, o totalio mais facilitador a escrita e leitura.

17 15. Comment on the reasons for your choice of the options in 14.
(Comente a volta dos motivos da sua escolha em 14)

Porque o mais importante é ter uma ortografia simples que facilite o processo de ensino e aprendizagem.

17 15. Comment on the reasons for your choice of the options in 14.
(Comente a volta dos motivos da sua escolha em 14)

Tanto a escrita/grafia Catlica quanto a Evangelica precisam de ser simplificadas nos seguintes aspectos.

17 15. Comment on the reasons for your choice of the options in 14.
(Comente a volta dos motivos da sua escolha em 14)

Porque assim lá existem o problema é só fazer deles mais simples.

17 15. Comment on the reasons for your choice of the options in 14.
(Comente a volta dos motivos da sua escolha em 14)

A preocupação tem como a simplificação, da escrita por os elementos serem simples para sua leitura e compreensão por facilitar a aprendizagem e a comunicação escrita.

Sample 2: Problems in teaching Umbundu

16. Comment on problem(s), if any, that you think there is with Umbundu language teaching and learning?
(Comente a volta de quaisquer problemas, se os houver, que achas que o ensino e aprendizagem do Umbundu enfrenta?)

Tem que ser que se aprendem como a ortografia. E precisam identificar este as partes para evitar confusão.

16. Comment on problem(s), if any, that you think there is with Umbundu language teaching and learning?
(Comente a volta de quaisquer problemas, se os houver, que achas que o ensino e aprendizagem do Umbundu enfrenta?)

O Ministério da Educação precisa editar manuais para servirem de base para o ensino da língua na região Ombundu.

16. Comment on problem(s), if any, that you think there is with Umbundu language teaching and learning?
(Comente a volta de quaisquer problemas, se os houver, que achas que o ensino e aprendizagem do Umbundu enfrenta?)

Além do conteúdo curricular escrito deveria haver o mesmo conteúdo no formato digital em PDF e em áudio ou vídeo e nos livros atuais há mais muita coisa através de filmes sobre história alguma importância delas para desenvolver a habilidade comunicativa de falar, ouvir e escrever porque quase não há material em áudio e vídeo.

THANK YOU SO MUCH FOR YOUR COOPERATION
MUITO OBRIGADO PELA COOPERAÇÃO
Brasá, 2020

18 16. Comment on problem(s), if any, that you think there is with Umbundu language teaching and learning?
(Comente a volta de quaisquer problemas, se os houver, que achas que o ensino e aprendizagem do Umbundu enfrenta?)

O principal é a falta de bibliografia e de literatura diversificada nesta língua.

18 16. Comment on problem(s), if any, that you think there is with Umbundu language teaching and learning?
(Comente a volta de quaisquer problemas, se os houver, que achas que o ensino e aprendizagem do Umbundu enfrenta?)

Tem que ser a experiência que vem a ter nos aulas de Umbundu. Atualmente não há. Há uma grande falta de materiais didáticos, fontes de pesquisa e materiais disponíveis na internet.

18 16. Comment on problem(s), if any, that you think there is with Umbundu language teaching and learning?
(Comente a volta de quaisquer problemas, se os houver, que achas que o ensino e aprendizagem do Umbundu enfrenta?)

O que falta a dizer é que a complexidade das duas variantes ortográficas tanto a católica como a protestante sempre existiram. Seria bom que houvesse uma harmonização ortográfica da escrita para um bom ensino desta língua.

Appendix 4: Written samples from the respondents: standard orthography tendencies

Sample 3: Catholic-like orthographical standard patterns

a. The baby knows it and knows the Lord. He is a witch. (O bebé sabe e conhece o Senhor. Ele é feiticeiro.)
Ngāngā o yīkūenda wakulihā ngāla. Eye Onganga.

b. The King has sprinkled the bag; the queen has sprinkled the outdoors. (O Rei borrifou o saco, a Rainha borrifou no vão.)
Soma wanhānhā ondjete, Nasoma wanhānhā posamua

c. The hippopotamus uses to defecate here. (O hipopótamo costuma defecar aqui.)
Ongeve yisole okunila palo/ ongeve yinilanila palo.

The snail slithers. (O caracol escorrega.)
Ehotio lilisielenena

Dad has left and I stayed yawning. (O pai partiu e eu fiquei a bocejar. Estou com raiva.)
Tate wandisia, ame ndasala lokuta onhānhā. Ngāsi lonhengō

Misery causes hunchback. (A miséria causa corcunda.)
Ongongo yikoka ongōngō.

a. The baby knows it and knows the Lord. He is a witch. (O bebé sabe e conhece o Senhor. Ele é feiticeiro.)
onganga yatulinbra e yatulinbrisa

b. The King has sprinkled the bag; the queen has sprinkled the outdoors. (O Rei borrifou o saco, a Rainha borrifou no vão.)
Osoma wanhānhā ondjete, Nasoma wanhānhā posamua

c. The hippopotamus uses to defecate here. (O hipopótamo costuma defecar aqui.)
Ongeve yatula okunila palo.

d. The snail slithers. (O caracol escorrega.)
Ehotio lilisielena

e. Dad has left and I stayed yawning. (O pai partiu e eu fiquei a bocejar. Estou com raiva.)
tate wandisa ame ndasala okutiso onhānhā

f. Misery causes hunchback. (A miséria causa corcunda.)
ofungui lenena ongōngō.

a. The baby knows it and knows the Lord. He is a witch. (O bebé sabe e conhece o Senhor. Ele é feiticeiro.)
O Ngāngā hi e waikulihā ngāla

b. The King has sprinkled the bag; the queen has sprinkled the outdoors. (O Rei borrifou o saco, a Rainha borrifou no vão.)
O soma wanhānhā ondjete Calye na soma wanhānhā posamua

c. The hippopotamus uses to defecate here. (O hipopótamo costuma defecar aqui.)
Ongeve yinilanila palo

d. The snail slithers. (O caracol escorrega.)
Ehotio lilisielena

e. Dad has left and I stayed yawning. (O pai partiu e eu fiquei a bocejar. Estou com raiva.)
tate wandisia calye ame ndasala

f. Misery causes hunchback. (A miséria causa corcunda.)
E fuquim enema

you are in it: (Tenta o teu máximo de escrever as seguintes frases na língua Umbundu para que se saiba o quanto podes):

a. The baby knows it and knows the Lord. He is a witch. (O bebé sabe e conhece o Senhor. Ele é feiticeiro.)
onganga yitehi kulenda yakulihā yigala

b. The King has sprinkled the bag; the queen has sprinkled the outdoors. (O Rei borrifou o saco, a Rainha borrifou no vão.)
Osoma wanhānhā ondjete, Nasoma wanhānhā posamua

c. The hippopotamus uses to defecate here. (O hipopótamo costuma defecar aqui.)
Ongeve yinyanya palo

d. The snail slithers. (O caracol escorrega.)
Ehotio lilisielena

e. Dad has left and I stayed yawning. (O pai partiu e eu fiquei a bocejar. Estou com raiva.)
Tate wandisia ndasala lokutwa onhānhā

f. Misery causes hunchback. (A miséria causa corcunda.)
uhukwi ukoka ongōngō

a. The baby knows it and knows the Lord. He is a witch. (O bebé sabe e conhece o Senhor. Ele é feiticeiro.)
Ngāngā o yīkūenda wakulihā ngāla. Eye Onganga.

b. The King has sprinkled the bag; the queen has sprinkled the outdoors. (O Rei borrifou o saco, a Rainha borrifou no vão.)
Soma wanhānhā ondjete, Nasoma wanhānhā posamua

c. The hippopotamus uses to defecate here. (O hipopótamo costuma defecar aqui.)
Ongeve yisole okunila palo/ ongeve yinilanila palo.

d. The snail slithers. (O caracol escorrega.)
Ehotio lilisielena

e. Dad has left and I stayed yawning. (O pai partiu e eu fiquei a bocejar. Estou com raiva.)
Tate wandisia, ame ndasala lokuta onhānhā. Ngāsi lonhengō

f. Misery causes hunchback. (A miséria causa corcunda.)
Ongongo yikoka ongōngō

Appendix 5: Written samples from the teacher/trainer questionnaire on the Umbundu orthography harmonization (continued)

Sample 4: Protestant-like orthographical standard patterns

a. The baby knows it and knows the Lord. He is a witch. (O bebe sabe e conhece o Senhor. Ele é feiticeiro.)
Omama yici kuenda ya kulinha nala
Eye omganga.

b. The King has sprinkled the bag; the queen has sprinkled the outdoors. (O Rei borrifou o saco, a Rainha borrifou no vão.)
Soma wabamba angete, Nasoma ubamba
elinava.

c. The hippopotamus uses to defecate here. (O hipopotamo costuma defecar aqui.)
Ngwene ysole okunila palo.

d. The snail slithers. (O caracol escorrega.)
Enhwa lichelela.

e. Dad has left and stayed yawning. (O pai partiu e eu fiquei a bocejar. Estou com raiva.)
Tate wanda ame machata lo kulata
enhamha waci komhena.

f. Misery causes hunchback. (A miséria causa corcunda.)
Ungongo yi hoka onono.

a. The baby knows it and knows the Lord. He is a witch. (O bebe sabe e conhece o Senhor. Ele é feiticeiro.)
O mama yici kuenda ya kulinha nala
Eye omganga.

b. The King has sprinkled the bag; the queen has sprinkled the outdoors. (O Rei borrifou o saco, a Rainha borrifou no vão.)
O soma ya nhamha angete, nasoma
enhamha posamua.

c. The hippopotamus uses to defecate here. (O hipopotamo costuma defecar aqui.)
Ngwene iniania palo.

d. The snail slithers. (O caracol escorrega.)
Ehotio lichelela.

e. Dad has left and stayed yawning. (O pai partiu e eu fiquei a bocejar. Estou com raiva.)
Tate wanda ame machata lo kulata
enhamha waci komhena.

f. Misery causes hunchback. (A miséria causa corcunda.)
Ungongo inena onono.
 Thank you very much for your cooperation

a. The baby knows it and knows the Lord. He is a witch. (O bebe sabe e conhece o Senhor. Ele é feiticeiro.)
Omama yici kuenda ya kulinha nala
Eye omganga.

b. The King has sprinkled the bag; the queen has sprinkled the outdoors. (O Rei borrifou o saco, a Rainha borrifou no vão.)
Soma wabamba wambaha angete,
Nasoma wambaha posamua.

c. The hippopotamus uses to defecate here. (O hipopotamo costuma defecar aqui.)
Ngwene ysole okunila palo.

d. The snail slithers. (O caracol escorrega.)
Ehotio lichelela.

e. Dad has left and stayed yawning. (O pai partiu e eu fiquei a bocejar. Estou com raiva.)
Tate wanda ame udsiala okuta okinava
waci komhena.

f. Misery causes hunchback. (A miséria causa corcunda.)
Ungongo yi hoka onono.

a. The baby knows it and knows the Lord. He is a witch. (O bebe sabe e conhece o Senhor. Ele é feiticeiro.)
Omama ya kulinha nala kuenda yi ci

b. The King has sprinkled the bag; the queen has sprinkled the outdoors. (O Rei borrifou o saco, a Rainha borrifou no vão.)
Soma wa nana nyanga konjeka, Naso-
ma wa nyanga posamua.

c. The hippopotamus uses to defecate here. (O hipopotamo costuma defecar aqui.)
Ngwene ya siata okunila palo.

d. The snail slithers. (O caracol escorrega.)
Ehotio li sielena.

e. Dad has left and stayed yawning. (O pai partiu e eu fiquei a bocejar. Estou com raiva.)
Tate wa euda ame udsiala okuta
okunyanwa namal si ci sole

f. Misery causes hunchback. (A miséria causa corcunda.)
Ungongo yi hoka onono!
Okukusini okukilokoka onono.
 Thank you very much for your cooperation

Appendix 6: Test of Umbundu (CSO) 1

PROVA TRIMESTRAL

Classe: 10 ^a A	Duração: 90 min.
Ano Lectivo 2021	Período: Tarde
II Trimestre	Data: 20.03.201
Disciplina: Língua Nacional	

Tanga l'evando oseteko yove noke tambulula l'unate apulilo osi vakusuñgamisila.

Walilongisa olonduko vy'ovikwata vyo vondjo y'elilongiso, kwenda esilivilo vikwete. Omo lyatcho, tambulula apulilo vakusuñgamisila.

- 1.1. Tukula ovikwata vyavelapo vali enene.
- 1.2. Popya nhe okwete okulinga otcho otchitumãlo tchove tchelilongiso katchikanholehe?
2. Kw'evyi vyatyamela k'ovinhama, tchikale evyi vyo vimbo ale vyo vusenge, vyosi vikasilili eposwiso ly'ofeka ndomo ove watchilikulihila.
 - 2.1. Esilivilo lipi likwete olongombe v'omwenho w'omunu ukwakulima?
 - 2.2. Ovinhama vipi vyo vimbo wakulihã okuti akamba amwele yavyo?
3. Pitiya olondaka evyi v'elimi ly'umbundu.
 - 3.1. Ondjo yetu y'elilongiso yaposoka omo yikwete ovikwata vyosi.
 - 3.2. Ositu y'ongulu yipepa vali hambì ositu y'ovinhama vikwavo.
 - 3.3. Pì tulimbukila omunu walya ositu y'ohombo? K'elemba!
 - 3.4. Otchimbungu otchinhama tchimwe tchivi.

Osande yiwa

Appendix 7: Test of Umbundu (CS0) 2

PROVA TRIMESTRAL

Classe: 11 ^o Período: Ano Lectivo: 2021/2022 II Trimestre Disciplina: L. Nacional	Duração: 90 min. Data: 01.04.2022
---	--------------------------------------

SÉRIE - A

LÊ ATENTAMENTE A SUA PROVA E COM MUITA CALMA RESPONDA AS QUESTÕES.

Ulonga

Eteke limwe, ukāyi lomola wahe, vanda kepya okukateya epungu kwenda okufela ekapa ale otchitina kumwe lotchipoke. Eci vakamālā kekumbi ombela yawa, ovo kavakwatele apa vavunda ombela.

Ukāyi lomola wahe vasokolola oukvunda vocipunda, eci vakevelelapo kamwe katito, ohosi kwenda omola wahe vangilavo vocipundo, kwendje yapa ukāyi lomōla wahe vakala lusumba walwa, yu valiylula oco ohosi yitile.

1. Sonehā tchiwa olondaka evi ndomo vakutchilongisa.

PROVA TRIMESTRAL

Classe: 11 ^o A, B, C e D Ano Lectivo: 2021/2022 II Trimestre Disciplina: Met. L. Nacional	Período: Tarde Duração: 90 min. Data: 01.04.2022
---	--

SÉRIE - A

LÊ ATENTAMENTE A SUA PROVA E COM MUITA CALMA RESPONDA AS QUESTÕES.

1. Popya esilivilo likwete ombela k'olongundja.
2. Tukulavo olosāyi ombela yiloka vali enene.

Boa Sorte

Appendix 8: Test of Umbundu (CS0) 3

PROVA TRIMESTRAL

Classe: 11^a Período:
Ano Lectivo: 2021/2022
II Trimestre
Disciplina: L. Nacional

Duração: 90 min
Data: 01.04.2022

SERIE - A

LÊ ATENTAMENTE A SUA PROVA E COM MUITA CALMA RESPONDA AS QUESTÕES.

Ulonga

Eteke limwe, ukäyi lomola wahe, vanda kepya okukateya epungu kwenda okufela ekapa ale otchitina kumwe lotchipoke. Eci vakamalã kekumbi ombela yawa, ovo kavakwatele apa vavunda ombela.

Ukäyi lomola wahe vasokolola oukvunda vocipunda, eci vakevelelapo kamwe katito, ohosi kwenda omola wahe vangilavo vocipundo, kwendje yapa ukäyi lomola wahe vakala lusumba waiwa, yu valiyula oco ohosi yitile.

1. Sonehä tchiwa olondaka evi ndomo vakutchilongisa.

Boa sorte

Appendix 9: Test of Umbundu (a combination of PSO, JWSO and the SSO) 4

Ano Lectivo: 2021 / 2022	IIº Trimestre
1ª Classe	Duração: 120'

PROVA TRIMESTRAL DE LÍNGUA NACIONAL UMBUMDU

Tanga Kuenda tambulula apulilo akuaimo:

1º Tokeka olondoka ndomo vilikuaĩ:

Mãe	Tate
Avó	Mayĩ
Pai	Pakulu

2º Telisa lovisonehuã (m, t, s, nd):

____ osi ____ a ____ u

epa ____ u

Appendix 10: Test of Umbundu (a combination of PSO, JWSO and SS0) 5

Ano Lectivo: 2021 /2022	IIº Trimestre
2ª Classe	Duração: 120'

PROVA TRIMESTRAL DE LÍNGUA NACIONALUMBUNDU

Tanga Kuenda tambulula apulilo akuaimo:

1- Telisa apangu lovisonehuã (o, i, t, u, e, nd):

- a) M_s_
- b) _ a _u

2- Sonehã atendelo a, velimi liofeka:

- a) 2 -
- b) 10 -
- c) 7 -
- d) 11 -

Appendix 11: Test of Umbundu (a combination of PSO, JWSO and the SSO) 6

Ano Lectivo:2021 /2022	IIº Trimestre
3ª Classe	Duração: 120´

PROVA TRIMESTRAL DE LÍNGUA NACIONALUMBUNDU

Tanga Kuenda tambulula apulilo akuaimo:

1º Pongolola ocipama eci lolondaka vikuava.

- a) Bom dia!
- b) Como estão?

2º Tokeka ulala olosaĩ:

Janeiro	Kupemba
Março	Susu
Maió	Kupupu

3º Sonehã velimi liofeka atendelo:

- a) 20 -
- b) 15 -
- c) 5 -
- d) 30 -
- e) 17 -
- f) 22-

Appendix 12: Test of Umbundu (a combination of PSO, JWSO and SS0) 7

Ano Lectivo: 2021 /2022

IIº Trimestre

4ª Classe

Duração: 120'

PROVA TRIMESTRAL DE LÍNGUA NACIONALUMBUNDU

Tanga Kuenda tambulula apulilo akuaimo:

1º Pongolola ocipama eci lolondaka vikuava.

- a) Bom dia!
- b) Como estão?

2º Tokeka ulala olosaĩ:

Janeiro	Kupemba
Março	Susu
Maió	Kupupu

3º Sonehã velimi liofeka atendelo:

- g) 20 -
- h) 15 -
- i) 5 -
- j) 30 -
- k) 17 -
- l) 22 -

Appendix 13: Test of Umbundu (SSO) 8

PROVA TRIMESTRAL DE UMBUNDU

Ano lectivo: 2021/22
I-Trimestre
Especialidade: O.E.P
Disciplina: Língua Nacional

11ª classe
Período: Manhã
Duração: 90 min.
Data: 10-12-2021

"Ukulihišo wavelapo ongusu" (O conhecimento supera a força)

Ondolo

Okofeko kayongwile okuputiwa esinga ,eci capitila p'epuluvu lyahé watumála kwenda upemuli wafetika okuteta esinga lyahé ,eci akamála okuteta esinga, okofeko kalimbuka akuti esinga lyahé lyapwamo.

Onepa ya tete

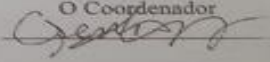
1-Tukula amanu vapanga onepa yolusapo.
2-Nye okofeko kakapangele kupemuli?
3-Nye calipita lesinga lyahé?
4-Eci upemuli apanga lokofeko cafina?

Onepa ya vali

1-Pitisa olondaka evi volulwo:
A)ongwambe b)ehondyo c)onjeke d)ocikutu e)okahombo
2-Kalye olondaka vyaco vitemisa kwasusululandaka.
3-Tukula olonduko vyomanu ava,tokekisa olondaka:
3.1-U otunga olonjo a) onjimbi
3.2-U wendisa ombinete b) onjali
3.3-U opanga ovopange yu kafetiwa c) usakuli
3.4-U osikila konjo yamále d) utungi
3.5-U wimba ovisungo e) uwendisi
3.6-U onena omóla kilu lyeve f) ukombe
3.7- U otata olombey g) ukwacali

4- **sonéha atendelo:**
a)Ise yange otambula(171.624) kolombongo.
5-Sonéha ciwa ulala wolondaka:
a)olosanji v'ocumbo conelého vikasi .
b)jyange okwete olosanji inanu vyalwa.
c)ilupukisa ombwa olosanji.
Upange uwa!

O Professor

O Coordenador


Appendix 14: Test of Umbundu (SSO) 9

Ano lectivo: 2020/21
Duração: 90 min.
Especialidade: O.E.P
II: Trimestre

Grupo: 1
Classe:
Período: Tarde
Data: ___/___/2021

Citekulu okwete alima ekwi la tatu kwanyamo, ise ya^{yahe}ke onalavayi yokombutika yesaku, kombaka
Oloneke vyosi leméle, ise yahe wenda kupange, wambata vocipete caye onariye omo kapondola ukutyuka konjo oku^{la}la. Ina yahe eye mbongela, osala k^{onjo} okutata vamanji yahe vatito vanda kosikola okutanga kekumbi, noke kwenda kocitanda okilandisa ombisi. Vonjo yavo vakasi omanu epandu, okukapa ko mapulu Mita. Citekulu wendenda vo kepya konene oko yolui okukwatisa ko mapulu yahe okulima.

1. Nôla ño etambululo lyaswapo. (escolha a resposta certa).

1.1. Tupula omanu vapanga onepa yepata lyaCitekulu?
a) Citekulu, katanya, Mita; b) Citekulu, ise, ina, vamanji, kwenda makulu yahe. c) Vilima, Mbongela, ise, Mita.

1.2. Alima añami akwete Citekulu?
a) Akwi atatu; b) Epandu c) ekwi la tatu.

1.3. Ise Yahe otalavaya pi?
a) Usakuli; b) Ulongisi; c) usonéhi

1.4. Ina yahe: a) otata olombey; b) otata ovinyama olandisa pocitanda;

1.5. Citekulu okwete vamanje vakulu?
a) Kakwete; b) okwete vavali; c) okwete umosi.

1.6. Monjo ya Citekulu mukwete vali omunu umuwe. Elye?
a) Inanu; b) Pakulu c) Makulu

2. Tukula eteke, osayi kwenda ulima okasi lokupanda oseteko.

Upange Uwa!

O Coordenador Docente

Appendix 15: Test of Umbundu (SSO) 10

ANOLECTIVO:2021/22 11ª classe

II-Trimestre Especialidade:O.E.P

Disciplina: Língua Nacional Duração:90 minutos

Data: 8-04-2022

“Ukulihisowavelapoongusu” (O conhecimento supera a força)

Prova Trimestral de Umbundu

Ondolo

UVEYI WO COVID-19

Uveyiwo Covid-19 usambuka lombili, owo wafetikila ko China, kosāyi ya Cembanima, kulima wolohulukāyi vivali lekwi leceya, noke walisandola kolofeka vyalwa mekonda lyomanu. Mulo mofeka yetu YONGOLA uveyi waco weya ko sāyi yelombo kulima wakwamako.

Omunu watete wakwatiwa luveyi owu osiñiwa kohupale lwo Luanda. O covid uveyi umwe ukasi lokumāla omwenyo womanu omo andi kaukwete esaku.Tuliteyili uveyi wo Covid lokusukula ovaka etu lovava kwenda onjapāwu.

Tanga ciwa oseteko yove noke tambulula lutate apulilo alikwāy (leia atentamente a prova e responda as questões que se seguem).

Onepa ya tete

1-O covid-19 nye?

2-Pi uveyi wo Covid wafetikila?

3-Tukula ulima kwenda osāyi uveyi owu wafetika?

a) Yipitisa kepulilo yelikalo.

b) Tukula onepa kwenda ekalo lyohundungulwa eyi:

“Ame ndakwatiwa kovo luweyi wo Covid-19”

2- Sokisa po olondaka lalumbambo (prefixos) asesamēla:

a) Ocindele ---ci ---kweteekovaliyela.

b) Omunu yu ---kwete olondunge.

c) Okamōlaaka ---kasilavi.

d) Olupasaolu ---nene.

3- Tukula eteke, osāyikwendaulimaokasilokupangaoseteko.

4- Sonēha ciwaulalawolondaka:

a) Olosanji v' ocumbo conelēho vikasi .

b) Yange okwete olosanji inanu vyalwa.

c) Ilupukisa ombwa olosanji.

Upange uwa!

A professor

COORDENADOR

Appendix 16: Test of Umbundu (SSO) 11

Docente:
Disciplina: L.N. II
Ano:1º - Di urno

Dura o da prova: 90 minutos
Data da Prova: 04/05/2022
Curso: L nguas

1. Soneh  olonduko vit lo vitukula ovinyama vyovusenge, okuti olo "prefixo" vyavyo vyatendelo a. (Escreva cinco substantivos que designam animais selvagens, cujos prefixos s o das seguintes classes).(5 valores)
- a) 5
 - b) 7
 - c) 9
 - d) 11
 - e) 12
1. Kaliye soneh  olohundunkulwa lokuvokiyakolonduko vyaco olondaka evi. Kwama ulala, ndomo ulekisiwa(Agora forme frases, acrescentado a esses nomes as seguintes palavras. Siga a ordem em que lhe s o apresentadas). (15 valores).
- a) Ocamanduko cokulekisa ca go, etendelo lyulala (11), ocinduke colusimbu, vonendelwa yoseteko yokulisoka. (Pronome demonstrativo simples, numeral ordinal (11) adjectivo primitivono grau comparativo de igualdade).
 - b) Ocamanduko cokulekisa ca go, etendelo lyulala (21), ocinduke casu gamel, vonendelwa yoseteko yokusulapo. (Pronome demonstrativo simples, numeral ordinal (21) adjectivo derivadono grau comparativo de inferioridade).
 - c) Ocamanduko cokulekisa ca go, etendelo lyulala (71), ocinduke colusimbu, vonendelwa yoseteko yokuvelapo. (Pronome demonstrativo simples, numeral ordinal (71) adjectivo derivadono grau comparativo de superiidade).
 - d) Ocamanduko cokulekisa ca go, etendelo lyulala (101), ocinduke casu gamela, vonendelwa yovonene volumwe wonengandaka. (Pronome demonstrativo simples, numeral ordinal (101) adjectivo derivadono grau superlativo absoluto anal tico).
 - e) Ocamanduko cokulekisa ca go, etendelo lyulala (252), ocinduke casu gamela, vonendelwa yovonene volumwe waleluka. (Pronome demonstrativo simples, numeral ordinal (252) adjectivo derivado no grau superlativo absoluto sint tico).

Appendix 17: Test of Umbundu (SSO) 12

Docente:
Disciplina: L.N. I
Ano:1º - Diurno

Duração da prova: 90 minutos
Data da Prova: 23/02/2022
Curso: Línguas

1. Sonehã olonduko vitálo vyocisoko cimosi, okuti olo “prefixo” vyavyo vyatendelo evi. (Escreva cinco substantivos monoclasses, cujos prefixos são das seguintes classes). **(5 valores)**
 - a) 3
 - b) 4
 - c) 6
 - d) 7
 - e) 11
1. Kaliye sonehã olohundunkulwa lokuvokiyakolonduko vyaco olondaka evi. Kwama ulala, ndomo vilekisiwa(Agora forme frases, acrescentado a esses nomes as seguintes palavras. Siga a ordem em que lhe são apresentadas). **(15 valores)**.
 - a) Ocamanduko coñgwate comunu watete, ocinduke casuñgamela, vonendelwa yoseteko yokulisoka. (Pronome possessivo da 1ª pessoa, adjectivo derivado, no grau comparativo de igualdade).
 - b) Ocamanduko coñgwate comunu wavali, ocinduke colusimbu (vonendelwa yoseteko yokuvelapo). (Pronome possessivo da 2ª pessoa, adjectivo primitivo, no grau comparativo de superioridade).
 - c) Ocamanduko coñgwate comunu watatu, ocinduke casuñgamela, vonendelwa yoseteko yokusulapo. (pronome possessivo da 3ª pessoa, adjectivo derivado, no grau comparativo de inferioridade).
 - d) Ocamanduko coñgwate comunu watete, ocinduke colusimbu vonendelwa yavelapovali yo “analítico”. (Pronome possessivo da 1ª pessoa, adjectivo primitivo, no grau superlativo absoluto analítico.)
 - e) Ocamanduko coñgwate comunu wavali, ocinduke casuñgamela, vonendelwa yavelapovali “sintético”. (Pronome possessivo da 2ª pessoa, adjectivo derivado, no grau absoluto sintético).

Appendix 18: Test of Umbundu (SSO) 13

PROVA PARCELAR DE LÍNGUA NACIONAL II

Ano lectivo 2021/2022

4º ano – Pós Laboral

Data:16. 12.2021

Duração: 90 minutos

1.Sonehá olonduko vitālo vyocisoko cimosi, okuti olonungilandaka vyavo vyovisoko evi (escreva cinco substantivos monoclasses, cujos prefixos são das seguintes classes): (5 valores)

- a) 3
- b) 5
- c) 7
- d) 9
- e) 11

2.Kaliye tunga olhundungulwa lokuvokiya kolonduko wasonehá ovyamanduko vyoŋigate kwenda ovinduke ndeci: (Agora forme frases, acrescentando aos substantivos escritos pronomes possessivos e adjectivos, seguindo o seguinte esquema): (7,5 valores)

a) Ocamanduko comunu watete kwenda ocinduke colusimbu (pronome possessivo da 1ª pessoa e adjectivo primitivo).

b) Ocamanduko comunu wavali kwenda ocinduke casuŋgamela (pronome possessivo da 2ª pessoa e adjectivo derivado).

c) Ocamanduko comunu watatu kwenda ocinduke colusimbu (pronome possessivo da 3ª pessoa e adjectivo primitivo).

d) Ocamanduko comunu wavali kwenda ocinduke casuŋgamela (pronome possessivo da 2ª pessoa e adjectivo derivado).

e) Ocamanduko comunu watete kwenda ocinduke colusimbu (pronome possessivo da 1ª pessoa e adjectivo primitivo).

3. Sonehá vali olhundungulwa vyokilu lokukapa ovinduke kolonendelwa evi: (Volte a escrever as frases acima, colocando os adjectivos nos seguintes graus):(7,5 valores)

a) Onendelwa yoseteko yokulisoka (Grau omparativo de igualdade).

b) Onendelwa yoseteko yokuvelapo (Grau comparativo de superioridade).

c) Onendelwa yoseteko yokusulapo (Grau comparativo de inferioridade).

d) Onendelwa yovonene volumwe wonengandaka (Grau superlativo absoluto analítico)

e) Onendelwa yovonene volumwe lwaleluka (Grau superlativo absoluto sintético).

Docente da Cadeira

O chefe do Departamento

Appendix 19: Analytical table of how sounds are coded in the various written works of the Umbundu orthography (author's data)

Sound	Example of corpus/source	Most probable influencing orthographical tradition	Adopting Religious Tradition		Author and publication date	Author's Academic Status
			Catholic	Protestant		
[ʃ]	Oxako (leaves)	Portuguese, e.g.: Oxalá!	No	No	Chimbinda (2015, pp. 30 – 31)	Catholic priest
	Isja (father of)	Umbundu, e.g. Isietu (from “ise” – father and “yetu” - our)	Yes	Yes	De Oliveira (2012, pp. 922 – 923)	Portuguese Researcher/Teacher
[g]	Songela (sharpen)	English, e.g.: g et, g ive, Kiswahili, e.g.: wa g eni; giza	Yes	Yes	SUZA, (2013, p. v) Palomé (1967, pp. 39; 42)	Coursebook designer Center for Applied Linguistics

	<i>“on̄eleya</i> (grave)	Kokota language, Santa Isabel Island, e.g.: gato- ḡonu	No	No	Jimbi & Sicala (2020, p. 130)	Researchers
	Katengue (Toponym)	Portuguese, e.g.: formigueiro	No	No	Queiroz (1978, p. 112)	Journalist
[ɲ]	ovo- <u>nh</u> a (feather)	Portuguese, e.g.: minha, vinho	Yes	No	Do Nascimento (1894, p. 2);	A medical doctor of the Royal Army
	Onyohā (snake)	Kiswahili, e.g.: nywi	No	Yes	SUZA (2013, p. 18)	Teachers
	K’eñulu (in the nose)	Spanish, e.g.: Spaña	No	No	De Oliveira (2012, p. 1029)	Researcher
[ŋ]	Waniṅga (in feces)	Umbundu, e.g. Owiṅgi	No	No	De Oliveira (2021, p. 1027)	Researcher
	Ñgo (sole)	Umbundu, e.g. viṅgami	Yes	No	António, Cuianda & Bonifácio (2012, p. 10)	Course designers
	Ng’o (sole)	Umbundu, e.g.: ong’oma	No	No	Pedro <i>et al.</i> (2013, p. 14)	Researchers
	Osuñila (is dozing)	Umbundu, e.g.: Ñala	No	Yes	IESA (2009, p. 217)	Clergy

[tʃ]	Ochikumbu, (mutilated) Ochili, (it's true)	English, e.g.: much, chamber	No	No	Do Nascimento (1894, p. 2)	Researcher
	Tjituñu (Ravine) Tjimbundu (anthroponym)	Otjiherero, e.g.: <i>Otjiingirisa</i>	No	No	The New Otjiherero Dictionary English Herero Otjiherero Otjiingirisa (2010) available at https://dicg5fhpqu3ab.cloudfront.net/gkzjbmsnhulvywfpqxqditooaac8c9bd3b8e6.pdf	A student of science
	Hatchoko (It's not)	English, e.g.: watch	Yes	No	Chimbinda (2015, p. 30) Malumbu (2007)	Catholic Priest, researcher
	Uc̣ito (birth)	Italian, e.g.: vicino	No	Yes	Mezzadri (2003, p. 60)	Course designer
[dʒ]	Onjevo (hunting)	English, e.g.: jack, judge	No	Yes	Do Nascimento (1894, p. 2) Hambly (1934, p. 237)	Medical Doctor African ethnologist
	Ondjali (progenitor)	English, e.g.: adjective	Yes	No	Chimbinda (2015, p. 30) Malumbu (2007)	Catholic Priest, Researcher

[s]	Osoma (King)	Kiswahili, e.g.: sambusa, sisi	Yes	Yes	SUZA (2013, p. v)	Course designers
	Saçoma (anthroponym)	Portuguese, e.g.: moça	No	No	Jimbi & Sicala (2020, p. 130)	Researchers
[k]	Ualunguka (smart)	Kiswahili, e.g.: Karibuni,	Yes	Yes	SUZA, 2013, p. 16	Course designers
	Çunene (big)	Portuguese, e.g.: Cabinda	No	No	Queiroz (1978, p. 93)	Journalist
	Caluquembe(Toponym)	Portuguese, e.g.: quando	No	No	Queiroz (1978, p. 93)	Journalist
[w]	Umue (somebody)	Portuguese, e.g.: nua, atuar	No	Yes	Do Nascimento (1894, p. 6)	
	Wanda (has gone)	English, e.g.: watch	Yes	Yes	Chimbinda (2015, p. 88)	Catholic priest, researcher
[j]	lene (yours)	Portuguese, higiene	No	Yes	Do Nascimento (1894, p. 7)	
	Yapa (meanwhile)	English, e.g.: Yes	Yes	Yes	Chimbinda (2015, p. 88)	Catholic priest, researcher

Appendix 20: The three most influential variations of the Umbundu orthography per source (7 phonemes and their orthographic representation)

Sound	Catholic orthography	Source	Protestant orthography	Source	Heterographical variations	Source
[tʃ]	<tch> e.g. “ot <u>ch</u> ili”	Embimbiliya Li kola (2017, p. 1058)	<c> e.g. “oc <u>i</u> li”	Hinário Evangélico (2009, p. 306)	<ch>; <tj> e.g. “och <u>i</u> sapa”; “ot <u>j</u> isapa”	Valente (1964, p. 157) Alves, (1951, I, p. 279)
[dʒ]	<Dje> e.g. “Kw <u>en</u> dje”	Embimbiliya Likola (2017, p. 1058)	<j> e.g. ““K <u>ue</u> nje”	Embimbiliya Li Kola (1963, p. 163)		
[ʃ]	<?>		<sie> e.g. “ <u>s</u> iel <u>ŷ</u> ha”	Embimbiliya Li Kola (1963, p. 163)	<x>; <sy> e.g. “ov <u>ŷ</u> ãĩ”; “wax <u>ŷ</u> ”; “ot <u>j</u> is <u>y</u> ãyi”	Chimbinda (2015, pp. 37-39) Alves, (1951, I, p. 924)
[k]	<k> e.g. “Kw <u>en</u> dje”	Embimbiliya Likola (2017, p. 1058)	<k> e.g. “k <u>ue</u> nje”	Embimbiliya Li Kola (1963, p. 163)		

[ŋ]	<ñg> e.g. “ñgo”	Embimbiliya Likola (2017, p. 1058)	<ñ> e.g. “ño-o”	Embimbiliya Li Kola (1963, p. 163)	<ḡ>; <ng’> e.g. “yalengela”; “ong’oma”	Alves, (1951, I, p. 748) Pedro et al. (2013, p. 14)
[ɲ]	<nh> e.g. “v’omwenho”	Embimbiliya Likola (2017, p. 1058)	<ny> e.g. “lomuenyo”	Embimbiliya Li Kola (1963, p. 159)	<ñ> e.g. “okatjiñama”	Alves, (1951, II, p. 1687)
[w]	<w> e.g. “Kwendje”	Embimbiliya Likola (2017, p. 1058)	<u> “Kuenje”	Embimbiliya Li Kola (1963, p. 163)		

Appendix 21: Religion variations of the Umbundu orthography seen from the students' translations (How respondents actually write in Umbundu)

Words	Umbundu pronunciation	Catholic respondents	Frequency of variations	Protestant (Evangelic) respondents	Frequency of variations
Baby	[ɔɾaɾa]	ngãngã; o ñgãñgã; nganga; o ngãnga; oñgãnga	5 times	oñaña; o ñaña; nãã; onãhã	4 times
Lord	[ɾaɾa]	Ngãla; ñgala; ngala	3 times	Ñala; nala; nãã	3 times
Witch	[ɔnganga]	Onganga; o nganga	2 times	Onganga	0 time
Sprinkled	[wanɾaɾa]	Wanhãhã; wanhanha; wanyanya	3 times	wañaña; wanhanha; wanhãhã; wanyanya	4 times
Uses to defecate	[jinjanja]	Inyanya; inhanha; iniania	3 times	Iniania; iniãniã; yinyanya	3 times
Slithers	[liʃelena/lisjelena]	Lilisielenena; lisyelena; lichelena; liatchelena	4 times	Lisielena; lichelena	2 times
Yawns	[ɔɾaɾa]	Onhanha; onyanya	2 times	Onhanha; Onhãhã; onyanya; onyãnyã	4 times
Misery	[ɔngongo]	Ongongo; o ngongo	2 times	Ongongo	0 time
Hunchback	[ɔɾɾɾ]	Ongõngõ; Ongõngo; ongongo; ongõngõ	4 times	Oñoño; onõnõ	2 times

Appendix 22: Chronological development of the Umbundu orthography of 4 lexemes + 3 expressions and their proponents

AUTHORS	DEVELOPMENTAL TRACES OF THE UMBUNDU ORTHOGRAPHY (in 110-Year time)						
English words	“hen”	“bat”	“rat”	“farms”	“dry season”	“small-pox”	“How much...?”
Koelle (1854)	“ósāṅḍṣi”	“erima”	“ómuku”	“owāḽā”	“ōkuénye”	“odzingōṅo”	“dṣin.ōami...?”
Bleek (1869)	“osandyi”	“erima”	“omuku”	“ovāḽya”	“okuenye”	“otyíngongo”	“tyiṅgami”
Capello & Ivens (1881)	“osandji” “osanji” “osan’gi”	“erima”	“omucu”	“ovāḽia”	“okuenhe”	“ot’chingongo” “otchingongo”	”T’chimgami?” “tchin’gami?”
Sanders and Fay (1885)	“osanje,”	“elima”	“omuku”	“ovapja”	“okwenye”	“Ocingongo”	“ciñame...?”
Torrend (1891)	“ossandyi,” “osandyi”	“elima”	“omuku”	“ovapja”	“okuenye”	“Ocingongo”	“chingami...?”
do Nascimento (1894)	“osanji,”	“elima”	“omuku”	“ovapja”	“okuenhe”,	“Ochingongo”	“chinhami...?”
Lecomte (1899)	“osangi,”	“elima”	“omuku”	“ovapja”	“okuenhe”	“Otchingongo”	“chiṅgami...?”
Hambly (1934)	“osanji” “osandji”	“elima”	“omuku”	“ovapja”	“okweñe” “okwenye”	“Ocingongo” “otcingongo”	“chiṅgami...?”
Alves (1964)	“osandji,”	“elima”	“omuku”	“ovapya”	“okweñe”	“Otjingongo”	“tjiṅgami...?”
Valente (1964)	“osanji”	“elima”	“omuku”	“ovapya”	“okwenhe”	“Ochingongo”	“chiḡami”

Appendix 23: Comparative table of the existing orthographical standards, the ameliorated proposal and the respective remarks on a more disjunctive approach

CSO	PSO	JWSO	Ameliorated Proposal NP/CA	Remark
<p>1V'eteke lyatcho, Yesu watunda v'imbo kwendje wakatumala k'ekulo lya kalunga.</p>	<p>1Keteke liaco Yesu wa tunda vonjo, yu (kuenje) wa tumala kohulo yokalunga.</p>	<p>1Keteke liaco, Yesu wa tunda konjo kuenje wa tumāla kohulo yokalunga</p>	<p>1Ke teke lia čo, Yesu watunda vo njo, yuu (kwenje) watumālā ko hulo yo kalunga.</p>	<p>It is preferable to take a disjunctive approach to the orthography for the fact that:</p> <p>a) In “v’eteke” and “keteke” one has (“vo”, “ko” ≈ in (the, a, an); “e” ≈ 3rd person of “be” and “teke”≈ day). As “k’e teke” is literally “in what is day”, meaning “in the day”, thus, easier to teach for Portuguese speakers.</p> <p>b) “lyatcho” and “liaco” have “lia”≈ “of the”, and “čo”≈so/this way. While “lyatcho” is excessively phonetic (omitting, in this way, the minimal units of meaning, i.e. the morphemes), “liaco”, like the former, appears as one word (making it difficult for teaching and learning).</p> <p>c) <č> is used to substitute the <tch> for the sake of economy (excessive three letters) and <c> for its confusing with the sound [k] as linked to the letters “a”, “o” and “u”.</p> <p>d) “kwendje” and “kuenje”. While the first has excessive letters for the sound[ɖʒ] <dj>, the latter takes <u> for the native sound [w] and takes <j> for the native sound [ɖʒ]. As <j> tends to be pronounced as [ʒ] for the majority of Portuguese speaking learners and so, eroding the original native Umbundu sound for it, this proposal suggests the use of <ĵ> as a differential, under the principles of economy/simplicity.</p>

watunda	wa tunda	wa tunda	watunda	The CSO “Watunda” has been here adopted because it eases the task of teaching the morphemes given that the morpheme “wa” in “tunda” indicates the 3 rd person singular. It would be easier to teach them conjunctively for a learner whose first language is Portuguese.
awiñgi	owiñi	owiñi	O wingĩ	The velar sound [ŋ] has demonstrated to be a problematic sound to write in Umbundu. In fact this problem has been discussed in Valente (1964) who believes that the tilde diacritic should be put above the letter <g> and not above the <n> as it currently is. The reason is that the tilde is used to nasalise the sounds that letters represent and, in this case the letter in question is <g> which is commonly used for the velar sound [g]. And this makes it right to have the tilde above <g> as proposed in the tables in this work.
kavalingi	Ka va lingi	Ka va lingi	ka valinḡi u panḡe	Here, the innovation has been done with the use of the <i>macron</i> diacritic which makes the <ḡ> permanently hard and pronounced solely as [g]. This ends with the confusion with the sound [ʒ] that the letter <g> represents when followed by <e> and <i> in Portuguese.
wakatumala	wa tumala	wa tumāla	watumālā	The innovation here resides in the nasalization of all possible vowels as Umbundu has proven to be one of the most nasalized Bantu languages.