

University of Stellenbosch
Faculty of Arts
Department of Afrikaans & Dutch
Program for Lexicography

The Standard Translation Dictionary as an Instrument in the Standardization of Fang

By
Thierry Afane Otsaga



**Dissertation presented for the Degree of
Doctor in Literature in Lexicography**

Promoter: **Prof Rufus H. Gouws**

April 2004

A decorative border with a repeating geometric pattern surrounds the central text.

Declaration

I, the undersigned, hereby declare that the work contained in this dissertation is my own original work and that I have not previously in its entirety or in part submitted it at any University for a degree.

Signature

Stellenbosch, April 2004

Summary/Opsomming

English

This dissertation offers metalexicographical criteria for the compilation of standard translation dictionaries in non-standardized languages. It particularly focuses on the role that the proposed dictionary model can play in the language standardization, with special reference to Fang. These criteria are based on different theoretical frameworks. Criteria for the dictionary structure are particularly based on Hausmann & Wiegand's General Theory of Lexicography, while a discussion of the role that dictionaries can play in the standardization process of languages is particularly based on criteria introduced by Zgusta. This dissertation is structured as follows:

- Chapter 1, *Introduction and problem statement*, presents the motivation for the choice of the present topic, as well as the theoretical frameworks that I use to build my metalexicographical criteria.
- Chapter 2, *Fang and its dialects*, focuses on the concerned language and its dialects. Apart from a historical overview of the Fang people, a wide inventory of works that have been implemented in Fang, as well as the choice and the motivation of the standard dialect are presented.
- Chapter 3, *Dictionaries and corpora*, highlights the strong relationship between the compilation of modern dictionaries and

electronic corpora. More precisely this chapter intends to demonstrate the importance of electronic corpora in the lexicographic practice. In this regard, the importance of some software and new corpus methods are also demonstrated.

- Chapter 4, *Aspects of the dictionary structure*, mainly focuses on different structures of the dictionary and how data should be spread and organized in each dictionary component. A detailed account is given of the structure of each dictionary component.
- Chapter 5, *Dictionaries and standardization*, focuses on the main purpose of the dissertation and demonstrates why and how the dictionary can be an instrument in the standardization process of languages, with specific reference to Fang.
- Chapter 6, *Concluding remarks*, reviews all chapters by highlighting the focal points of each of them. Some perspectives or potential new developments are foreseen in order to pave the way for the elaboration of new theoretical frameworks and the improvement of the proposed dictionary model.

Afrikaans

Hierdie proefskrif bied metaleksikografiese kriteria vir die samestelling van standaard vertalende woordeboeke vir nie-gestandaardiseerde tale. Die fokus is veral op die rol wat die voorgestelde woordeboekmodel kan speel in taalstandaardisering, met spesifieke verwysing van Fang. Hierdie kriteria is gebaseer op verskillende teoretiese raamwerke. Kriteria vir die woordeboekstruktuur is veral gebaseer op Hausmann en Wiegand se algemene leksikografieteorie, terwyl 'n bespreking van die rol van woordeboeke in die standaardiseringsproses van tale veral gebaseer is op kriteria wat deur Zgusta voorgestel is. Die proefskrif is soos volg saamgestel:

- Hoofstuk 1 bied 'n motivering vir die keuse van die betrokke onderwerp asook vir die teoretiese raamwerke waarvolgens die metaleksikografiese kriteria geformuleer is.
- Hoofstuk 2 fokus op die betrokke taal, Fang, en sy dialekte. Naas 'n historiese oorsig van die Fang sprekers word 'n wye keuse uit die Fang literatuur asook die keuse van 'n standaarddialek aan die orde gestel.
- Hoofstuk 3 wys op die sterk verhouding tussen die samestelling van 'n moderne woordeboek en elektroniese korpora. Hierdie hoofstuk benadruk die belang van elektroniese korpora vir die leksikografiese praktyk. In hierdie verband word daar ook verwys na die belang van sekere sagtewareprodukte asook nuwe korpusmetodes.

- Hoofstuk 4 is veral gerig op die verskillende woordeboekstrukture en op hoe data versprei en in die verskillende woordeboekkomponente aangebied moet word. Die struktuur van elke woordeboekkomponent word in besonderhede bespreek.
- Hoofstuk 5 fokus op die hoofdoel van die proefskrif en wys hoe en waarom 'n woordeboek, met spesifieke verwysing na die situasie van Fang, 'n instrument kan wees in die taalstandaardiseringsproses.
- Hoofstuk 6 gee 'n oorsig oor die voorafgaande hoofstukke en wys op sekere belangrike aspekte wat bespreek is. Nuwe ontwikkelinge wat voorsien word om die weg te baan vir die ontwerp van nuwe teoretiese raamwerke ter verbetering van die voorgestelde model word beklemtoon.



Dedication

To my grand parents, who always showed how much they loved me before they past away. I am sure they still look after me from where they are. Angoué Ntoutoume, Mengue-Me-Ndong & Assengone Eyaga I dedicate this doctoral dissertation to you.

Table of contents

| | |
|--|--------------|
| Declaration | ii |
| Summary/opsomming | iii |
| Dedication | vii |
| Table of contents | viii |
| Acknowledgements | xiii |
| | |
| Chapter 1: Introduction and Problem statement | 1-50 |
| 1.1. Introduction | 1 |
| 1.2. Problem statement | 9 |
| 1.2.1. Introduction | 9 |
| 1.2.2. The Gabonese language situation | 9 |
| 1.2.3. Focus | 11 |
| 1.2.4. The source language | 12 |
| 1.2.5. The target group | 18 |
| 1.2.5.1. Characteristics of the target group | 18 |
| 1.2.6. Typological model | 22 |
| 1.2.6.1. Introduction | 22 |
| 1.2.6.2. Scerba's classification | 23 |
| 1.2.6.3. Malkiel's classification | 24 |
| 1.2.6.4. Al-Kasimi's classification | 25 |
| 1.2.6.5. Zgusta's classification | 28 |
| 1.2.6.6. Hartmann and James' classification | 31 |
| 1.2.6.7. Béjoint's classification | 33 |
| 1.2.6.8. Remarks | 35 |
| 1.2.6.9. Classification of bilingual dictionaries: a new vision | 36 |
| 1.2.6.10. The proposed dictionary model: A typological hybrid | 38 |
| 1.2.6.11. Standard features | 41 |
| 1.2.7. Dictionary functions | 42 |
| 1.2.7.1. General remarks | 42 |
| 1.2.7.2. Functions of the proposed dictionary model | 44 |
| 1.2.8. Survey of some aspects of some theoretical models and approaches | 45 |
| 1.2.8.1. Introduction | 45 |
| 1.2.8.2. Landau and the material collection | 46 |
| 1.2.8.3. Hausmann and Wiegand and dictionary components | 46 |
| 1.2.8.4. Gouws and the outer texts of the dictionary | 48 |
| 1.2.8.5. Zgusta and language standardization | 49 |
| 1.2.8.6. Gallardo and the role of the dictionary in the language standardization process | 49 |
| | |
| Chapter 2: Fang and its dialects | 51-86 |
| 2.1. Preliminaries | 51 |
| 2.2. Language and dialect | 54 |
| 2.3. Fang dialects in Gabon | 57 |
| 2.3.1. Introduction | 57 |
| 2.3.2. The Fang-Atsi dialect | 57 |
| 2.3.3. The Fang-Mekè dialect | 60 |
| 2.3.4. The Fang-Mvaï dialect | 63 |
| 2.3.5. The Fang-Ntumu dialect | 64 |
| 2.3.6. Fang-Okak dialect | 67 |
| 2.3.7. Fang-Nzaman dialect | 69 |
| 2.3.8. General works in Fang | 69 |
| 2.4. Choice of the Standard dialect | 70 |

| | |
|---|----------------|
| 2.4.1. General remarks | 70 |
| 2.4.2. Criteria | 71 |
| 2.4.3. Comments | 77 |
| 2.4.4. The proposed standard dialect | 83 |
| 2.5. Concluding remarks | 84 |
| Chapter 3: Dictionaries and corpora | 87-125 |
| 3.1. Introduction | 87 |
| 3.2. Corpora typology | 90 |
| 3.3. The material | 94 |
| 3.3.1. Introduction | 94 |
| 3.3.2. Material acquisition | 94 |
| 3.3.2.1. General approaches | 94 |
| 3.3.2.2. Criteria for the material selection | 96 |
| 3.3.2.3. Primary sources | 97 |
| 3.3.2.4. Secondary sources | 103 |
| 3.3.2.5. Tertiary sources | 104 |
| 3.3.2.6. Some remarks about text collection | 107 |
| 3.3.3. Material preparation | 110 |
| 3.3.3.1. Introduction | 110 |
| 3.3.3.2. Preparation of written material | 110 |
| 3.3.3.3. Text encoding | 112 |
| 3.3.2.4. Corpus query tools | 115 |
| 3.3.2.5. Comments | 118 |
| 3.3.2.6. Corpus updating | 119 |
| 3.3.2.7. Corpora and standard dictionaries | 121 |
| 3.3.2.8. Concluding remarks | 123 |
| Chapter 4: Aspects of dictionary structure | 126-217 |
| 4.1. Introduction | 126 |
| 4.2. Data distribution structure | 127 |
| 4.2.1. General remarks | 127 |
| 4.2.2. Data distribution structure in the central list | 129 |
| 4.2.2.1. Introduction | 129 |
| 4.2.2.2. Macrostructure | 130 |
| 4.2.2.2.1. General remarks | 130 |
| 4.2.2.2.2. Macrostructure in Fang dictionaries | 131 |
| 4.2.2.2.2.1. Introduction | 131 |
| 4.2.2.2.2.2. Macrostructure in the <i>Encyclopédie Pahouine</i> of Largeau | 131 |
| 4.2.2.2.2.3. Macrostructure in the <i>Lexique Fān-Français</i> of Martrou | 132 |
| 4.2.2.2.2.4. Macrostructure in the <i>Dictionnaire fang-français et français-fang</i> of Galley | 133 |
| 4.2.2.2.3. Different types of macrostructures | 134 |
| 4.2.2.2.4. Proposed macrostructure | 140 |
| 4.2.2.2.5. Selection of lemmata | 142 |
| 4.2.2.2.6. Category of the selected lemmata | 149 |
| 4.2.2.2.7. Form of lemmata | 150 |
| 4.2.2.2.8. Functions of lemma signs | 151 |
| 4.2.2.2.8.1. Introduction | 151 |
| 4.2.2.2.8.2. Data on spelling | 151 |
| 4.2.2.2.8.3. Lemma as key to the dictionary | 152 |
| 4.2.2.2.9. Linguistic units not to be selected as lemmata | 152 |
| 4.2.2.2.9.1. Morphemes | 153 |
| 4.2.2.2.10. Concluding remarks | 153 |
| 4.2.2.3. Microstructure | 154 |
| 4.2.2.3.1. General remarks | 154 |
| 4.2.2.3.2. Microstructure in existing Fang dictionaries | 156 |
| 4.2.2.3.2.1. Introduction | 156 |

| | | |
|--|--|----------------|
| 4.2.2.3.2.2. | Microstructure of the <i>Encyclopédie Pahouine</i> of Largeau | 156 |
| 4.2.2.3.2.3. | Microstructure of the <i>Lexique fân-français</i> of Martrou | 156 |
| 4.2.2.3.2.4. | Microstructure of the <i>Dictionnaire Fang-Français et français-fang</i> of Galley | 159 |
| 4.2.2.3.3. | Different types of microstructures | 161 |
| 4.2.2.3.3.1. | General remarks | 161 |
| 4.2.2.3.3.2. | On the distinction comment on form and comment on semantics | 162 |
| 4.2.2.3.3.3. | Unintegrated microstructure | 164 |
| 4.2.2.3.3.4. | Integrated microstructure | 167 |
| 4.2.2.3.3.5. | Semi-integrated microstructure | 170 |
| 4.2.2.3.3.6. | Rudimentary microstructure | 173 |
| 4.2.2.3.4. | Proposed microstructure | 173 |
| 4.2.2.3.4.1. | General remarks | 173 |
| 4.2.2.3.4.2. | Micro-architecture | 177 |
| 4.2.2.3.5. | Detailed description of microstructural features | 178 |
| 4.2.2.3.5.1. | Introduction | 178 |
| 4.2.2.3.5.2. | Equivalent relations in bilingual dictionaries | 181 |
| 4.2.2.3.5.2.1. | General remarks | 181 |
| 4.2.2.3.5.2.2. | Absolute equivalence | 181 |
| 4.2.2.3.5.2.3. | Partial equivalence | 183 |
| 4.2.2.3.5.2.4. | Surrogate equivalence | 186 |
| 4.2.2.3.5.2.5. | Remarks on translation | 188 |
| 4.2.2.3.5.3. | Addressing structure | 190 |
| 4.2.2.3.5.3.1. | General remarks | 190 |
| 4.2.2.3.5.3.2. | Lematic addressing | 191 |
| 4.2.2.3.5.3.3. | Sublemmatic addressing | 192 |
| 4.2.2.3.5.3.4. | Full lemmatic addressing | 192 |
| 4.2.2.3.5.3.5. | Non-lemmatic addressing | 193 |
| 4.2.2.3.5.3.6. | Remarks on addressing structure | 194 |
| 4.2.2.3.6. | Microstructure and cultural data | 195 |
| 4.2.3. | Data distribution structure in outer texts | 197 |
| 4.2.3.1. | Introduction | 197 |
| 4.2.3.2. | Front matter texts | 197 |
| 4.2.3.2.1. | General remarks | 197 |
| 4.2.3.2.2. | Front matter and user's guidelines | 199 |
| 4.2.3.3. | Back matter texts | 200 |
| 4.2.3.3.1. | General remarks | 200 |
| 4.2.3.3.2. | Back matter and complex articles | 202 |
| 4.3. | Other aspects of the dictionary structure | 205 |
| 4.3.1. | Access structure | 205 |
| 4.3.1.1. | General remarks | 205 |
| 4.3.1.2. | Outer access structure | 206 |
| 4.3.1.2.1. | Introduction | 206 |
| 4.3.1.2.2. | Remarks on outer access structure | 211 |
| 4.3.1.2.3. | Proposed outer access structure | 212 |
| 4.3.1.3. | Inner access structure | 212 |
| 4.3.2. | Mediostructure | 214 |
| 4.4. | Concluding remarks | 216 |
| Chapter 5: Dictionaries and standardization | | 218-280 |
| 5.1. | Introduction | 218 |
| 5.2. | Standardization: a better understanding of the concept | 219 |
| 5.3. | Why standardize a language? | 221 |
| 5.4. | Standardization and the writing system | 223 |
| 5.4.1. | Introduction | 223 |
| 5.4.2. | History of the alphabet | 223 |
| 5.4.2.1. | Introduction | 223 |
| 5.4.2.2. | Origins of the alphabet | 225 |
| 5.4.2.3. | Birth of vowels | 226 |

| | | |
|--------------------------------------|---|----------------|
| 5.4.2.4. | The wider international influence of the Greek alphabet | 226 |
| 5.4.3. | The writing of African languages | 228 |
| 5.4.3.1. | A brief overview | 228 |
| 5.4.3.2. | Alphabet and orthography | 229 |
| 5.4.3.2.1. | Definitions | 229 |
| 5.4.4. | Writing of Gabonese languages | 231 |
| 5.4.4.1. | Introduction | 231 |
| 5.4.4.2. | The Raponda Walker's alphabet | 233 |
| 5.4.4.3. | The April 1999 alphabet | 234 |
| 5.4.4.4. | The Rapidolangue's alphabet | 235 |
| 5.4.4.5. | Quality of the standard alphabet | 236 |
| 5.4.4.6. | Choice of the standard alphabet | 240 |
| 5.4.4.7. | Matter of tones in the writing of Gabonese languages | 243 |
| 5.4.4.7.1. | Introduction | 243 |
| 5.4.4.7.2. | Types of tones in Gabonese languages | 244 |
| 5.4.4.7.3. | Pros and cons of the transcription of tones | 245 |
| 5.4.4.8. | Conjunctivism or disjunctivism | 252 |
| 5.4.5. | Concluding remarks | 255 |
| 5.5. | Dictionaries and the standardization of Fang | 256 |
| 5.5.1. | Introduction | 256 |
| 5.5.2. | Role and Authority of the dictionary | 257 |
| 5.5.2.1. | Introduction | 257 |
| 5.5.2.2. | Influence of dictionaries in the speech community | 258 |
| 5.5.2.3. | Dictionaries and learners of a new language | 260 |
| 5.5.2.4. | Dictionaries and educational system | 260 |
| 5.5.2.5. | Dictionaries and socio-professional fields | 262 |
| 5.5.2.6. | Concluding remarks | 262 |
| 5.5.3. | Dictionaries and the development of Fang | 263 |
| 5.5.3.1. | Introduction | 263 |
| 5.5.3.2. | Dictionaries and contemporary Fang | 264 |
| 5.5.3.3. | Dictionaries and the writing and reading of Fang | 265 |
| 5.5.3.4. | Dictionaries and Fang as an academic subject | 266 |
| 5.5.3.5. | Dictionaries and Fang as language of education | 268 |
| 5.5.3.6. | Dictionaries and the development of the standard form of Fang | 269 |
| 5.6. | Standardization and electronic dictionaries | 275 |
| 5.6.1. | Introduction | 275 |
| 5.6.2. | Types of electronic dictionaries | 276 |
| 5.6.3. | Advantages of electronic dictionaries | 276 |
| 5.6.4. | Internet and the development of Fang | 277 |
| 5.6.4.1. | Introduction | 277 |
| 5.6.4.2. | Role of Internet in the standardization of Fang | 278 |
| 5.7. | Concluding remarks | 278 |
| Chapter 6: Concluding remarks | | 281-324 |
| 6.1. | Introduction | 281 |
| 6.2. | Introduction and Problem statement | 283 |
| 6.2.1. | Review | 283 |
| 6.2.2. | Perspectives | 285 |
| 6.3. | Fang and its dialects | 287 |
| 6.3.1. | Review | 287 |
| 6.3.2. | Perspectives | 289 |
| 6.4. | Dictionaries and corpora | 293 |
| 6.4.1. | Review | 293 |
| 6.4.2. | Perspectives | 296 |
| 6.5. | Aspects of dictionary structure | 298 |
| 6.5.1. | Review | 298 |
| 6.5.2. | Perspectives | 308 |
| 6.6. | Dictionaries and standardization | 310 |

| | |
|-------------------------|------------|
| Table of Contents | xvii |
| <hr/> | |
| 6.6.1. Review | 310 |
| 6.6.2. Perspectives | 321 |
| 6.7. General conclusion | 323 |
| Bibliography | 325 |
| Index | 334 |

Acknowledgements

The presentation of a doctoral dissertation is often the outcome of a long and difficult research process. To come up with good quality results, this process needs not only the effort and hard work of the researcher, but also the support and aid from people related to the him and/or his research. That is why I would like to express my deepest gratitude to:

- Prof. Rufus H. Gouws, my supervisor, for his academic guidance and support throughout the years. This work could not be completed without his expertise and availability.
- Prof. James D. Emejulu, the coordinator of the convention between the Gabonese Universities and the University of Stellenbosch, for encouraging me to come in this University and do my postgraduate studies in lexicography.
- My big brothers, Anicet Oyane Akaga and Alain Evah Nzé, for encouraging me to continue my studies. Their material and financial support have been of a great important for the completion of my studies.
- My big sisters, Vivette Assame Otsaga and Marie Edzidzi Otsaga, for always being there when I need them. Their good advices have been the light that illuminated the itinerary of my studies.

-
- My father, Robert Otsaga Afane, and my mother, Suzanne Ezzo Angoué, for teaching me that knowledge is power.
 - My girlfriend, Laetitia Christelle Obame N’na, for her support and understanding despite the distance. Her love has been of a great assistance during my down moments.
 - The Gabonese Government, through the Bourses & Stages, for sponsoring my studies.
 - All the International Office staff members from the University of Stellenbosch for making my stay in Stellenbosch easier.
 - The Buro Van Die WAT for guiding my first steps in lexicography and giving me the first practical training in the field.

I am also grateful to my brother and friend Dr. Paul Achille Mavoungou for reading this work and assisting me in some aspects of its realization.

Finally, I thank all the Gabonese students at the University of Stellenbosch, particularly my fellow students in lexicography, as well as any people who, directly or indirectly, have contributed to the success of this work. God bless all of them.

CHAPTER 1:

Introduction and problem statement

The theory of lexicography is not a theory merely for the sake of the theory. It is a practice-directed theory with as primary objective the formulation of a model to improve the quality of dictionaries. However, in order to achieve an improvement in the quality of the lexicographic practice, resulting in the compilation of better dictionaries, lexicographers have to be made aware of the existence of lexicographic theory and of the necessity to compile their dictionaries in accordance with criteria presented in such a theory (Gouws, 2001: 59).

1.1. Introduction

The standard translation dictionary as an instrument in the standardization of Fang is the title of this research project. Dictionaries differ from one another in several ways, including their aims, their scope and the subject(s) they cover. These differences coincide with the typological variation within the broader category of dictionaries. Translation dictionaries, which can be either bilingual or multilingual (Al Kasimi, 1977), are generally opposed to defining dictionaries (e.g. monolingual dictionaries). However, both translation and defining dictionaries are classified as linguistic dictionaries (Gouws, 2000). Linguistic dictionaries can be either diachronic or synchronic. Diachronic dictionaries are primarily concerned with the history, and the development of words (lexical units), in respect of both form and meaning. Whereas the task of diachronic dictionaries is to deal with the development of the lexicon, the purpose of synchronic dictionaries is to deal with the lexical stock of a language at one stage of its development (Al Kasimi, 1977:202).

The dictionary model I am proposing will mainly integrate synchronic features, even if some diachronic aspects could also be included. The term *translation dictionary* is often used as a synonym of the *bilingual dictionary*. In a way this is correct since most translation dictionaries are bilingual. However, a bilingual dictionary may not necessarily be designed for the purposes of translation; it may be specifically aimed to aid the dictionary user in learning a non-native language. In such a case, the dictionary will provide detailed explanations of lexical meanings of the object language in the learner's native language. The term *translation dictionary* should be used exclusively with reference to linguistic dictionaries that are only intended to provide interlingual translation equivalents and generally serve the purpose of translation (Burkhanov, 1998:248; as cited by Tarp 2002b).

In this work, the term *translation dictionary* will primarily refer to bilingual dictionaries, and the proposed model will be set to provide interlingual translation equivalents between French and Fang, primarily aimed at the translation from French to Fang.

By *standard dictionary*, I mean a dictionary presenting and treating the language in such a way that it can serve as a basis, an example or a principle to which Fang speakers (should) conform, or by which they are judged in their language use. This dictionary will primarily endeavour to reflect the standard variety of the target language (Fang). It will be normative in its approach. The norm can be conceived as the part of the total possibilities offered by the system which is considered "good", i.e., functionally most adequate (Zgusta, 1971). Standard dictionaries are the linguistic reference sources most commonly used by the average member of a speech

community and these users rely on the dictionary as an authoritative source of linguistic information. Where a language does not have an established collection of dictionaries, the typological pattern will be to compile one or more standard dictionaries. This type of dictionary plays an important role in the standardization process. Where a language is not yet fully standardized, it is customary to compile a standard translation dictionary before compiling a standard descriptive dictionary (Gouws, 2000:15).

The functions of a Fang standard dictionary could be formulated according to those described by Garvin concerning a standard national language (Zgusta, 1971:185):

- A unifying function; several dialect areas are united into a single standard language-community.
- A separatist function; a speech community is set off as separate from its neighbours.
- A prestige function and a frame-of-reference function; the prestige resulting from the possession of a standard language, and the function to serve as frame of reference for correctness and for the perception and the evolution of poetic speech.

Apart from any eventual feelings which different linguistic problems may excite in different areas of the world, the communicational needs of a modern society or a developing one, suffice to make the standard national language a means of communication much more desirable than particular dialects. Therefore, the development of a standard national form will be considered, in the majority of cases, a highly desirable task in such a language community (Zgusta,

1971:185). Dictionaries definitely influence users in their language use. Because of this influence, they become the sources of reference that speakers confer with for the best use of languages. Zgusta (1989:75-76) stressed three basic ways in which dictionaries can influence users:

- Selection of information: for instance, if from among several synonyms or other variants only one or two are listed in the dictionary, it can be presumed that the user who tries to communicate in the new standard language would use the listed rather than the non-listed variants (why would he otherwise consult, if he felt that he knew better?). The selection or non-selection of information is an influential factor, because users do check the dictionary, particularly when drafting written texts, and they frequently accept what the dictionary says.
- Labels and symbols that give information about status and register (*colloquial, informal, dialectal, obsolete, etc.*) of expressions, or that give indications of syntactic patterns also influence users.
- Another means by which larger dictionaries influence users are short commentaries of various types within the article or attached to it. Indeed, dictionaries with a historical slant can and do give a historically founded argument for or against a usage. Dictionaries of technical terminology sometimes give encyclopaedic arguments to motivate the selection of a term instead of a competing synonym. Some dictionaries frequently use "usage notes" that select the preferred form from competing syntactic, collocational or other variants, or warn

that using a word or expression involves some problems of which the user may not be aware. These usage notes inform the user about various possibilities and leave the final choice to him. The reader thus gets more involved and may develop a sensitivity and judgement of his own.

All those arguments show that a dictionary could really play an important role in the use of a language. The standard bilingual dictionary I propose could influence the future use of Fang. This dictionary will be an *instrument*, a vehicle in the standardisation process of Fang. I will focus on several aspects of translation dictionaries with emphasis on the bilingual dictionary with Fang as one member of the language pair.

The first part of this research will present the problem statement. In this section the nature of the issue will be specified, taking into account the meaning of the central concept posed in its problem statement. I shall give the reasons that guided me in the choice of Fang, what are the needs of the Fang community and why I chose the standard translation dictionary as an instrument in the standardization process of Fang.

The second step will be to present the different dialects of Fang. Mention will be made of the studies already undertaken in Fang and its dialects and the results thereof. Other academic background information relevant to Fang and its dialects will also be mentioned. This part will give me the opportunity to explain in detail why I am proposing Fang-Ntumu as the standard dialect. The criteria for the choice of the standard dialect will be discussed with the intention to provide this choice as objectively as possible.

The third step of this dissertation will deal with the relation between dictionaries and the compilation of electronic corpora. Data collection is an important part in the process of dictionary compilation. A language material collection policy is a set of guidelines for the collection of language material. In this dissertation, the main goal of the chapter regarding the database will be to review some practical stages in the process of corpus building, from the selection of sources through to mark-up, and assigning annotations to assembled texts. Mention will also be made of the population group from which the corpus data (oral & written) will come, with the objective to compile a standard French/Fang bilingual dictionary. Specific mention will also be made of the general preparation phase, the material acquisition phase and, the material preparation phase:

- The general preparation phase will deal with the foundation for the structure, contents and presentation of the final product. This stage will particularly be concerned with the identification, establishment, nature, extent and description of a dictionary basis, which suits the relevant dictionary project in the best possible way (Gouws, 2001:68). A dictionary basis can be described as the total of the source language material for the specific lexicographic process (Wiegand, 1998:139 as cited by Gouws, 2001: 68).
- The material acquisition phase will focus on the gathering of speech material from sources earmarked for the dictionary basis. A result of the material acquisition phase is the compilation of the lexicographic corpus (Gouws, 2001:69).

- The material preparation phase will concern the preparation of the collected material for the processing. In the case of oral material for instance, the recordings have to be transcribed and scanned into the computer for eventual inclusion in the corpus.

The fourth chapter will deal with some aspects of the dictionary structure. Hausmann and Wiegand (1989) have proposed a theoretical model for the component parts and structures of monolingual dictionaries. This model, which can also be used for translation dictionaries, will guide me to find a way in which the data distribution could be compiled. For the theoretical model I am proposing, the focus is on **the macrostructure, the microstructure, the access structure, the addressing structure, the mediostructure** and **the outer texts** of a translation dictionary. During the discussion on each of these topics, several examples will be given to show how the model could be applied to Fang.

Like most Bantu languages, Fang is a language with supra segmental elements (tone, accent, intonation, etc.), which play an important role in the functioning of this language. Amongst all these supra segmental elements, tones have the most visible and influential role (Afane Otsaga, 2000: 155). As such, tones¹ should be taken into account in the lexicographical treatment of Fang. Indeed, in Fang, one word can have several meanings according to the way it is pronounced. Although until now, in the majority of Fang dictionaries, this tone phenomenon has not been taken into

¹As said by Martinet A. (Baylon, C. & B. Fabre, 1990) tones are oppositions of voice pitches, which in certain languages are used as distinctive units.

account. In this research, I will propose some approaches that could help with the treatment of tones in Fang. These approaches could be applied to any language with the phenomenon of tones. The issue of tone will be discussed in the sections on standardization and writing.

The fifth and penultimate chapter, the main focus of this doctoral dissertation, has the purpose to identify the relationship between dictionaries and the language standardisation process. In fact, according to Gallardo (1980) & Zgusta (1989) dictionaries can influence language standardisation in several ways. Those ways will be discussed and analysed with particular reference to Fang. Specific attention will be paid to the spelling system, the standardisation process and the promotional policy. The electronic dictionary, whether on CD-ROM, online, or hand-held, will supersede the paper dictionary in ways unimaginable in the paper-dictionary dimension, just as the computer has completely superseded the typewriter (Prinsloo, 2001:139). This aspect should be taken into account, even in African languages, and that is why in this section, reference will also be made to the role that the new technology facilities could play in the development of African languages in general and Gabonese languages in particular, and how these facilities should be effectively used.

The sixth and last chapter of this dissertation will be the conclusion. Here, each chapter will be summarized and new perspectives will be emphasised. I will outline new theories and methods, and suggest improvement of the proposed model for the future. In other words, a flash back underlining the focal points of each chapter will be highlighted and some potential new

developments will be foreseen in order to pave the way for the elaboration of new theoretical frameworks and the improvement of the proposed dictionary model.

1.2. Problem statement

1.2.1. Introduction

The aim of this research is to devise a theoretical model for the compilation of standard dictionaries directed at the standardisation of the Gabonese languages. This research could also be applied to other multilingual environments by assisting with the lexicographic presentation of emerging standard languages. Fang will be the Gabonese language that I will use as illustration for the formulation of this model. The present chapter will focus on: (1) the Gabonese language situation, (2) the location of Fang varieties in Gabon, (3) the motivation for the choice of French as source language of the dictionary, (4) the presentation of the target group and the motivation for the choice of this group, (5) the presentation of features that should be included in the typological model (6) and finally, the outline of some aspects of the theoretical models I will use in this research project.

1.2.2. The Gabonese language situation

Many inventories of the Gabonese languages have been made e.g. by M. Guthrie (1948, 1953 and 1971), A. Raponda Walter (1960), A. Jacquot (1978) and J.T. Kwenzi Mikala (1987 and 1997). Kwenzi Mikala's inventories are the most recent and include more languages than the others. In the last classification (1997), Kwenzi Mikala has inventoried 62 Gabonese language forms that are grouped in 10 language units. The language forms are grouped

according to their mutual comprehension and the name of each language unit is from the expression "I say that", which serves to open a conversation in Gabonese languages (Kwenzi Mikala, 1990) as presented in the following table.

Table 1: Gabonese language forms classification according to Kwenzi Mikala (1997)

| Language Unit | Language forms |
|-------------------------|--|
| 1. Mazona | 1. Fang-Atsi 2. Fang-Mekè 3. Fang-Mvaï 4. Fang-Ntumu 5. Fang-Nzaman 6. Fang-Okak |
| 2. Myénè | 1. Myénè-Enenga 2. Myénè-Galwa 3. Myénè-Mpongwe 4. Myénè-Nkomi 5. Myénè-Orungu 6. Myénè-Okoa |
| 3. Mekena-Menaa | 1. Akélé 2. Ungom 3. Lisigu 4. Mbagwé 5. Metombolo 6. Seki 7. Tumbidi 8. Shaké 9. Wumpfu 10. Lendambama |
| 4. Mekona-Mangoté | 1. Ikota 2. Benga 3. Shamayi 4. Makongwé 5. Ndasa 6. Bakola |
| 5. Membé (Okandé-tsogo) | 1. Gétso 2. Gépinzi 3. Kandé 4. Gébobé 5. Gehimbaka 6. Gebiya 7. Ebongwe 8. Kota-Kota |
| 6. Méryé | 1. Gésira 2. Gibarama 3. Gibungu 4. Yipunu 5. Yilumbu 6. Yisangu 7. Ngubi 8. Civili 9. Yirimba 10. Yigama |
| 7. Métyé | 1. Yinzébi 2. Yitsengi 3. Yiwélé 4. Yibili 5. Liduma 6. Liwandzi 7. Yibongo |
| 8. Membéré | 1. Lembaama 2. Lékanini 3. Lindumu 4. Latégé 5. Latsitsége |
| 9. Mekena | 1. Bekwil 2. Shiwa 3. Mwèsa |
| 10. Baka | 1. Baka |

Table source: T. Afane Otsaga, 1999. *Esquisse phonologique et Morphologique du meka (langue Bantoue du Gabon de sigle A83)*. Rapport de Recherche en vue de l'obtention du grade de licence ès lettres, option Linguistique Africaine. Université Omar Bongo (UOB). Libreville.

In this inventory, Kwenzi Mikala includes languages and dialects. At the present stage of the research on Gabonese languages only two language units can, with certainty, be called a language. In fact, Fang-Atsi, Fang-Mekè, Fang-Mvaï, Fang-Ntumu, Fang-Nzaman and Fang-Okak are recognised as Fang dialects that Kwenzi Mikala calls

the Mazona language unit. Just like Myénè-Enenga, Myénè-Galwa, Myénè-Mpongwé, Myénè-Nkomi, Myénè-Orungu and Myénè-Okoa are accepted as dialects of the Myénè language unit. However, this distinction cannot be made for the other language units. Because it is still difficult to make a clear distinction between dialects and languages, it is not easy to say exactly how many languages there are in Gabon. Despite this difficulty, it is evident that Gabon is a multilingual country with several languages. Those languages have a wide demographic distribution throughout the country. This wide distribution is also the main reason for regional varieties. The new policy of the Gabonese government is to make provision for the eventual inclusion of local languages in the educational system in the near future. This goal will be properly achieved if the languages are standardized. In fact, all languages with dialects have faced the same problem, and one of the policies they used to solve this problem was to standardize the language. Gallardo (1980) has stressed the importance and the functional role of bilingual dictionaries as instruments in the standardization process of languages. He tries to show how dictionaries and the standardization process are closely linked. Fang, which is one the Gabonese languages whose dialects have been clearly identified, can be used as a good model in this regard.

1.2.3. Focus

Fang is the language that will be the focus of this dissertation. Spoken in four provinces in Gabon (Estuaire, Moyen-Ogooue, Ogooue-Ivindo and Woleu-Ntem) and in neighbouring countries (Cameroon, Equatorial Guinea, Sao Tomé and the Republic of the Congo), Fang increasingly displays differences between its dialects.

This growing variation between dialects tends to decrease the success of their mutual comprehension. This work will essentially focus its attention on Fang spoken in Gabon. In Gabon, Fang speakers constitute 30 percent of the population (Encarta Encyclopaedia, 2000) and populate 16 regions of that country (cf. table 2).

Table 2: Fang dialects locations

| Fang-dialects | Locations |
|---------------|--|
| Fang-Atsi | Lambaréné - Ndjolé - Bifoun |
| Fang-Mekè | Libreville - Kango - Ntoun - Foulenzem |
| Fang-Mvaï | Minvoul |
| Fang-Ntumu | Oyem - Bitam |
| Fang-Nzaman | Makokou - Koumameyong - Booué - Ovan |
| Fang-Okak | Cocobeach -Mitzic - Medouneu |

Table source: "Carte des parlers du Gabon". 1990, in Revue Gabonaise des Sciences de l'homme. N 2. LUTO. UOB. Libreville.

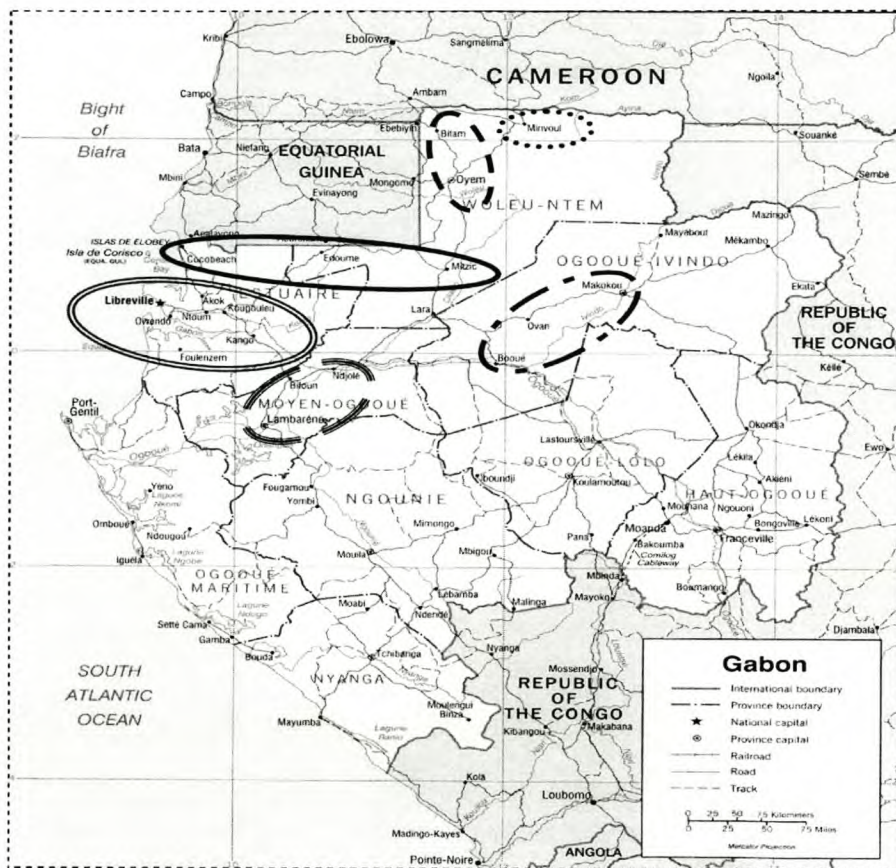
N.B.: The distribution of Fang dialects throughout Gabon can be seen on the map on the next page.

1.2.4. The source language

In this research project a model will be devised for a standard translation dictionary with French and Fang as the language pair. It will be a model for a *monoscopal* dictionary with French as source language and Fang as target language. The term *monoscopal* is used in the same way as in Hausmann and Werner (Gouws, 2001). According to them, scope refers to the language direction. A bilingual dictionary can be monoscopal with one range or direction (from language A to language B) or biscopal with two ranges or directions (from language A to language B and from language B to language A). When a monoscopal dictionary (A > B) is compiled in a systematic way, it could be easy for the lexicographer to make another dictionary containing the second scope (B > A) thanks to

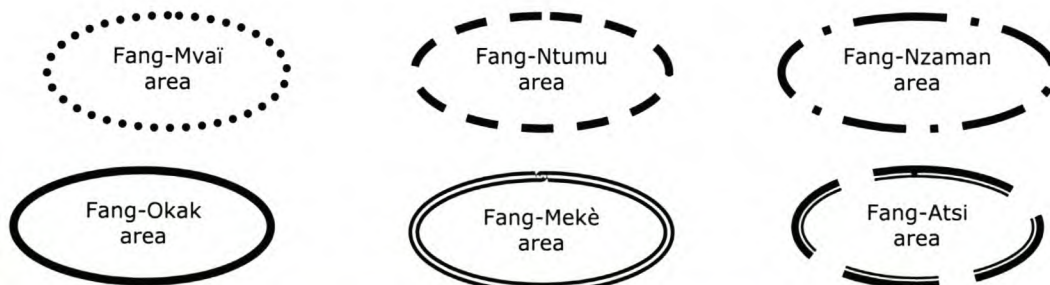
computer-assistance facilities. The compilation of a monoscopal dictionary is the ideal situation. This implies that for a given language pair that polyfunctional dictionary should have two volumes and each volume should be monoscopal (Hausmann as cited by Gouws, 2001:80).

Map 1: Fang dialects distribution in Gabon



Source Map: <http://www.lib.utexas.edu/maps/gabon.html>

Map special setting: Thierry Afane Otsaga



The way Hausmann & Werner use the terms *monoscopal* and *biscopal*, should not be confused with *mono-* and *bidirectional*. *Mono-* and *bidirectional* according to them refer to the targeted dictionary users² (Gouws, 2001:80).

The proposal to use French as the source language is motivated by the fact that the majority of Fang people in Gabon (particularly young people) communicate in French, even at home. Many young Fang speakers have an underdeveloped knowledge of the language of their parents. They need to be directed back to their language via French, the only official language of Gabon. The lexicographer, when compiling any dictionary, should take such user needs into account. These needs should always influence the genuine purpose of the dictionary. The French position in Gabon can be explained by a brief historical overview of the linguistic policy in this country. According to Kwenzi Mikala (1990), during the French colonisation the linguistic policy was regulated by two prescriptions:

- The first prescription, which can be traced back to 1539, from Villiers Cotterets, prohibited the use of languages other than French in all official administrative functions.
- The second prescription (in 1922) regulated teaching. One of its dispositions recommended that general teaching must be done exclusively in French. Only religion could be taught in local languages.

This linguistic policy has existed for a long time in Gabon even after independence. For instance, up to the 1980s it was prohibited in

²According to Hausmann and Werner's point of view, a dictionary is monodirectional when it is targeting only one category of users (mother tongue speakers of the source language for instance) and bidirectional when it is targeting two categories of users (mother tongue speakers of the source and the target language). However, in the common usage mono- and bidirectional are identical to the

schools for pupils to communicate in their mother tongue even during the break time. Anyone caught could be punished in several ways. The most popular punishment was to give them something heavy or silly to carry around their neck (with a rope) for a day or until they catch someone else speaking a local language. French still has a privileged position in Gabon, resulting from its status as sole official language, stipulated in the *Gabonese* constitution in article 2, title 1: *La Republique Gabonaise adopte le Français comme langue officielle*³. French is used as the exclusive language of public authorities (Politics and Economics) and in important Institutions. It is the language of writing, e.g. in educational systems; newspapers; correspondence; relations between enterprises or administrations and mastery of the modern technology (Moussirou Mouyama as cited by Mekui Missang, 1998). Because French dominates the intellectual life of the Fang people and controls several important activities, it can be considered as the language of everyday life. French also ensures communication between Gabonese people from different linguistic groups (Mekui Missang, 1998). All these examples explain the strong position of French today in Gabon.

I have carried out an investigation in Gabon between January and March 2001 in Minvoul, Oyem, Libreville, Ntoum and Kango. Young Fang speakers were the target group of this investigation (between 14 and 35 years old) and the analysis of the answers to the questions posed revealed the following results:

- **Question 1:** *Apart From Fang, what others languages do you speak?*

definitions they give for mono- and biscopal.

³ The Republic of Gabon adopts French as official language.

French: 100%
Another Gabonese language: 20%
English: 10%
Spanish: 5%
Other: 1%

- **Question 2:** *In which language do you do think, talk or dream most often?*

French: 70%
Fang: 29%
Other: 1%

- **Question 3:** *Which language do you think you speak very well?*

French: 65%
Fang: 34%
Other: 1%

- **Question 4:** *In which language do you communicate with your family members?*

French: 60%
Fang: 40%

- **Question 5:** *Which language do you use preferentially in the following fields?*

- **Academic field?**

French: 85%
Fang: 10%
Others: 5%

- **Work field**

French: 80%
Fang: 18%
Others: 2%

- **Urban field?**

French: 70%
Fang: 25%
Other: 5%

- **Rural field?**

Fang; 65%
French: 34%
Others: 1%

- **Question 6:** *What percentage of the day do you estimate using the following languages?*

French: 60%
Fang: 30%
Another Gabonese language: 7%
A foreign language: 3%

- **Question 7:** *Which language do you use to communicate with the following people?*

- Your father and your mother?

Fang: 60%
French: 39%
Others: 1%

- Your children or/and your nephews?

French: 80%
Fang: 19%
Others: 1%

- Your friends (also Fang)?

French: 70%
Fang: 29%
Others: 1%

- Elderly Fang people?

Fang: 95%
French: 4%
Others: 1%

- Young Fang people?

French: 70%
Fang: 25%
Others: 5%

- **Question 8:** *Can you read and write in Fang?*

No: 94%

Yes: 5%
A bit: 1%

The results show the strength of French among young Fang Gabonese people. That is why I propose French as source language for the compilation of a French/Fang bilingual dictionary. It is easier for someone to retrieve information in a dictionary from a language he knows very well. The need of young Fang speakers is to be able to transfer their knowledge from French to Fang.

1.2.5. The target group

1.2.5.1. Characteristics of the target group

The planning of any specific dictionary must always be related with the corresponding or expected group of users and their characteristics. A group of intended users of a specific dictionary must be determined in order to detect the specific needs of each specific type of user and decide on the corresponding data to be included in the dictionary. According to Tarp (2002), the following user characteristics must be taken into account when planning any dictionary:

1. Which language is their mother tongue?
2. At what level do they dominate their mother tongue?
3. At what level do they dominate a foreign language?
4. What is the level of their general cultural and encyclopaedic knowledge?
5. At what level do they dominate the special subject field in question?
6. At what level do they dominate the corresponding LSP in their mother tongue?

7. At what level do they dominate the corresponding LSP in the foreign language?

For the target group of the dictionary model I am proposing, the following answers could be given for each question pointed out by Tarp:

1. According to my investigation, a large part of the young Fang speakers do have French as their first language and most of them are more competent in French than Fang, even those ones who have Fang as their mother tongue. However, other members of the target group have Fang as their mother tongue and can also claim some competence in that language. For that reason, both French and Fang can be considered as the mother tongue of the target group.
2. Even if French and Fang can play the role of mother tongue for the target group, they still need to improve their knowledge in both languages. Therefore, inclusion of important data in both French and Fang will be very helpful for them (e.g. meaning paraphrases of lemmata and translation equivalents).
3. If both French and Fang are considered as mother tongues of the target group, it also means they are considered as foreign languages to them. Indeed, for the part of the target group whose mother tongue is French, Fang will be considered as foreign language. Conversely, the part of the target group that has Fang as its mother tongue will have French as foreign language. However, the group who do have Fang as its mother tongue have very limited competence in that language. Only a few of them can read and write in Fang. Their competence in Fang is mainly oral.

4. The general culture of the target group must be determined according to each educational level included in that group. In fact, the target group involves students from different levels from primary school to university (including the college and the high school). Ideally, a specific investigation should be carried out to determine the general culture of each level of the target group. The needs, in terms of knowledge, will be different according to the level of study for each potential dictionary user. When for instance high school students will need general and basic knowledge in any specific domain, university students will need detailed information in their domain of study.
5. The proposed dictionary will not be an encyclopaedia as such, but it will include many encyclopaedic features regarding specific fields that the target group is dealing with so that the user can have as much information as possible in special fields.
6. The part of the target group that has Fang as its mother tongue cannot read and write in Fang (in most cases). Therefore, it is difficult for the users from that category to really dominate any LSP in their mother tongue.
7. The part of the target group that has French as its mother tongue are in a better position. They can read and write in French, as that is the sole official language in Gabon. The target group deals with French almost in all their activities and different levels. Consequently they dominate their LSP better in French.

According to Hartmann and James (1998: viii), one of the principal advances in lexicography in recent years has been the focus on the user perspective, that is the realisation that different users have different reasons for using a dictionary, and that the dictionary can,

and should, respond to these. They add that whilst dictionaries have usually tried to satisfy the overall perceived needs of large classes of users; there has often been little attention paid to the specific needs of smaller groups or individuals. Even if a dictionary can be used by anybody, a dictionary project must have a specific target user group.

Dubois & Dubois (1971: 10), corroborate that point of view by saying that a dictionary plan must try to answer the questions of the target group. The lexicographer should take into account the needs of this target group. For the typological model I am going to present, the young Fang speakers constitute the main target group the dictionary is intended for. The following reasons guided me to choose young Fang speakers as the target group for a standard bilingual French/Fang dictionary:

- In the near future, local languages will be included in the Gabonese educational system. Gabonese languages will be taught in the same way as French, English, Spanish, etc, and will be, in some cases, the medium of instruction. In this regard, a translation French/Fang dictionary will be very helpful to young Fang students.
- According to my investigation, young Fang speakers are more competent in French than in Fang. Even though Fang may not be an academic subject for them, they still need a translation dictionary to communicate more easily in Fang (e.g. with old Fang speakers who do not know French).

For these reasons young Fang speakers form an important group that could be targeted by a lexicographer compiling a French/Fang translation dictionary. The dictionary model I am proposing will also

be bi-directional (as the term is used by Hausmann & Werner, 1991). In fact, apart from young Fang speakers, any French speaking person who wants to learn Fang (e.g. all francophone people, even Frenchmen) could use this dictionary. One of the goals of this dictionary is to include current terms related to the modern human activities; particularly those linked to science and technology (e.g. data processing, computer, internet, cell phone, mathematics, physics, chemistry, biology, etc.). Hence the dictionary model I propose will be useful to any Fang speaker or any teacher using Fang (as a subject or as a medium of instruction). Teachers will use it not only to help students, but also to improve their own competence in Fang.

1.2.6. Typological model

1.2.6.1. Introduction

Many researchers have made different dictionary classifications. Indeed, dictionaries have been classified according to several criteria, e.g. the size (small, medium or big), the purpose (descriptive, comprehensive, prescriptive or normative), the range (in alphabetical or thematic order), the treated languages (monolingual, bilingual or multilingual), the perspective (synchronic or diachronic), the scope (monoscopal or biscopal), the contents (linguistic or extra-linguistic), the users or the direction (for native and/or foreign speakers), the field (general or specific), etc. In a lexicographic study, it is important to be well aware of the fact that no single dictionary can be everything to everyone. There is an extended dictionary family and one has to realise where each dictionary fits into the typological structure and according to which criteria a dictionary has to be evaluated when determining its

typological status (Gouws, 2000:10). Before I discuss the type of dictionary I am going to propose in this project, an overview of some dictionary classifications will be made. These classifications will allow me to retrieve some features that I will propose for my dictionary model. I will particularly outline dictionary typologies proposed by Scerba (1940), Malkiel (1967), Al-Kasimi (1977), Zgusta (1971), Hartmann and James (1998), and Béjoint (2000).

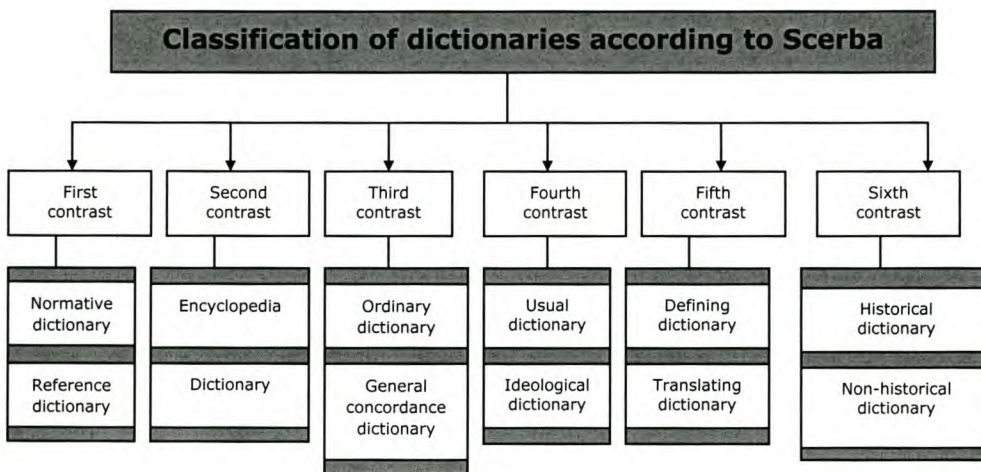
1.2.6.2. Scerba's classification

The Russian linguist L.V. Scerba made one of the earliest typological studies of dictionaries published in 1940 (Al Kasimi, 1977:12). Scerba's classification is based on the structural characteristics of possible dictionary types and he distinguishes six categories of contrast:

1. A normative dictionary, which dictates norms vs. a reference dictionary, which adopts a descriptive approach.
2. An encyclopaedia vs. a dictionary. The contrast here is based on the function of proper names in a language. Scerba thinks that proper names are part of the language and should be excluded from the dictionary.
3. An ordinary dictionary (e.g. defining or translating dictionary) vs. general concordance in which all the words are listed along with all quotations that can be found in texts, as is the case in a concordance of a dead language.
4. A usual dictionary (such as a defining dictionary or a translating dictionary) vs. an ideological dictionary that groups ideas or subjects.
5. A defining dictionary (e.g. a monolingual dictionary) vs. a translation dictionary (e.g. bilingual or multilingual dictionary).

6. An historical dictionary vs. non-historical dictionary. The purpose of the historical dictionary is to give all meanings of all words that belong to, and have belonged to a given language since its beginning.

Scerba's classification can be schematized as follows.



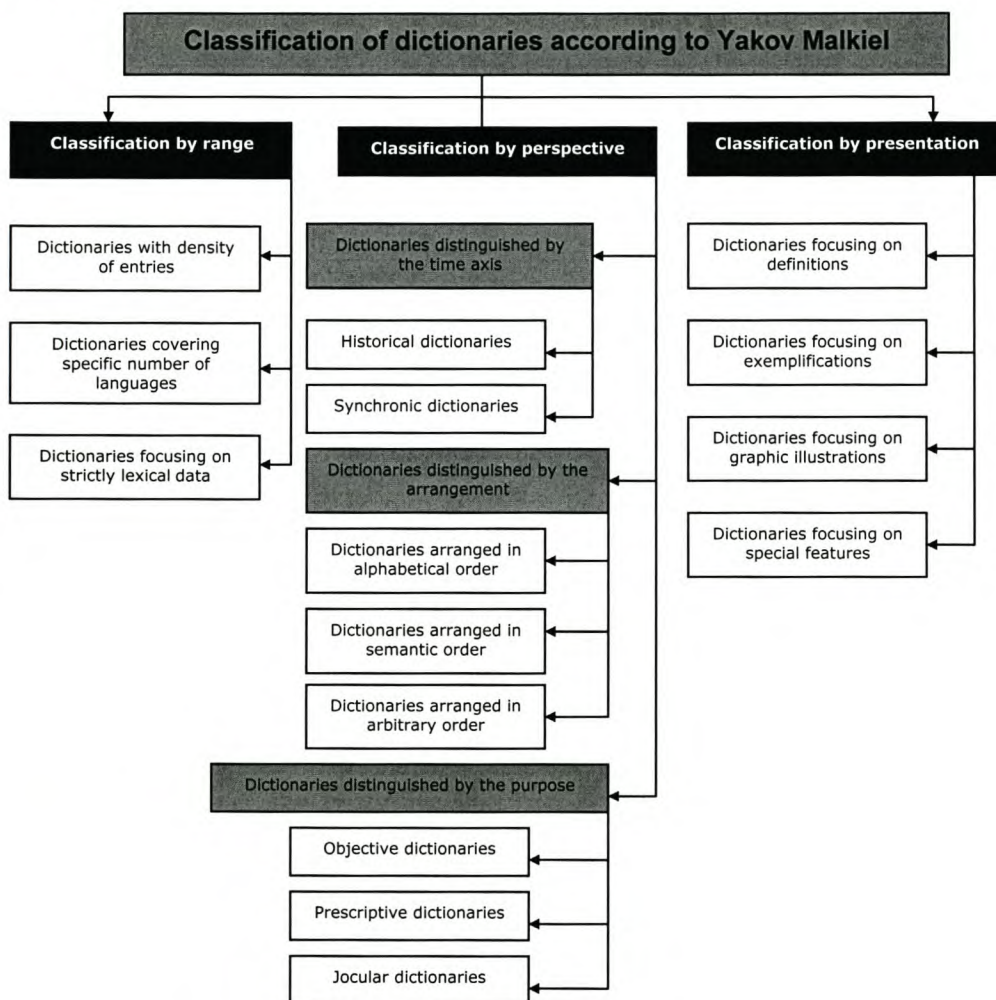
1.2.6.3. Malkiel's classification

In *Problems of lexicography* edited by Householder and Saporta (1967), Yakov Malkiel proposed a typological classification of dictionaries based on distinctive features. Malkiel retained three main criteria according to which dictionaries should be distinguished: the range, the perspective, and the presentation.

- The *classification by range*, which Malkiel considers as the most obvious and objective criterion, focuses on (a) density of entries, (b) number of languages covered or resorted to, (c) degree of concentration on strictly lexical data.
- The classification by perspective takes into account (a) the time axis, e.g. historical (dynamic) or synchronic (static); (b) the arrangement, e.g. alphabetical, semantic or arbitrary; (c) the purpose, e.g. objective, preceptive, or jocular.

- The classification by presentation focuses on four main aspects (a) the definition; (b) the exemplification; (c) graphic illustrations (including maps); (d) special features (e.g. location, pronunciation, etc.).

Malkiel’s classification can be schematised as follow.



1.2.6.4. Al-Kasimi’s classification

Al-Kasimi (1997) is one of the few lexicographers who made a classification of dictionaries by focusing on a specific type of dictionary. Although, he took into account the previous general classifications, Al-Kasimi’s classification only focuses on bilingual

dictionaries. He distinguished dictionaries according to their source, their scope or their purpose, and has set up seven contrasts of types of bilingual dictionaries:

1. Dictionaries for the speakers of the source languages vs. dictionaries for speakers of the target language. By the source language he means the language of the entries (lemmas), and the target one is the language of the translation or equivalents.

2. Dictionaries of the literary language vs. dictionaries of the spoken language. Modern linguists have made it clear that speech is the fundamental form of language activity, and writing is just a representation of that speech.

3. Dictionaries for production vs. dictionaries for comprehension. The distinction between a dictionary for production and one for comprehension is exhibited in the very first stages of making the dictionary (i.e. the choice of source and target languages). Another difference between the two dictionaries lies in the content and the structure of entries.

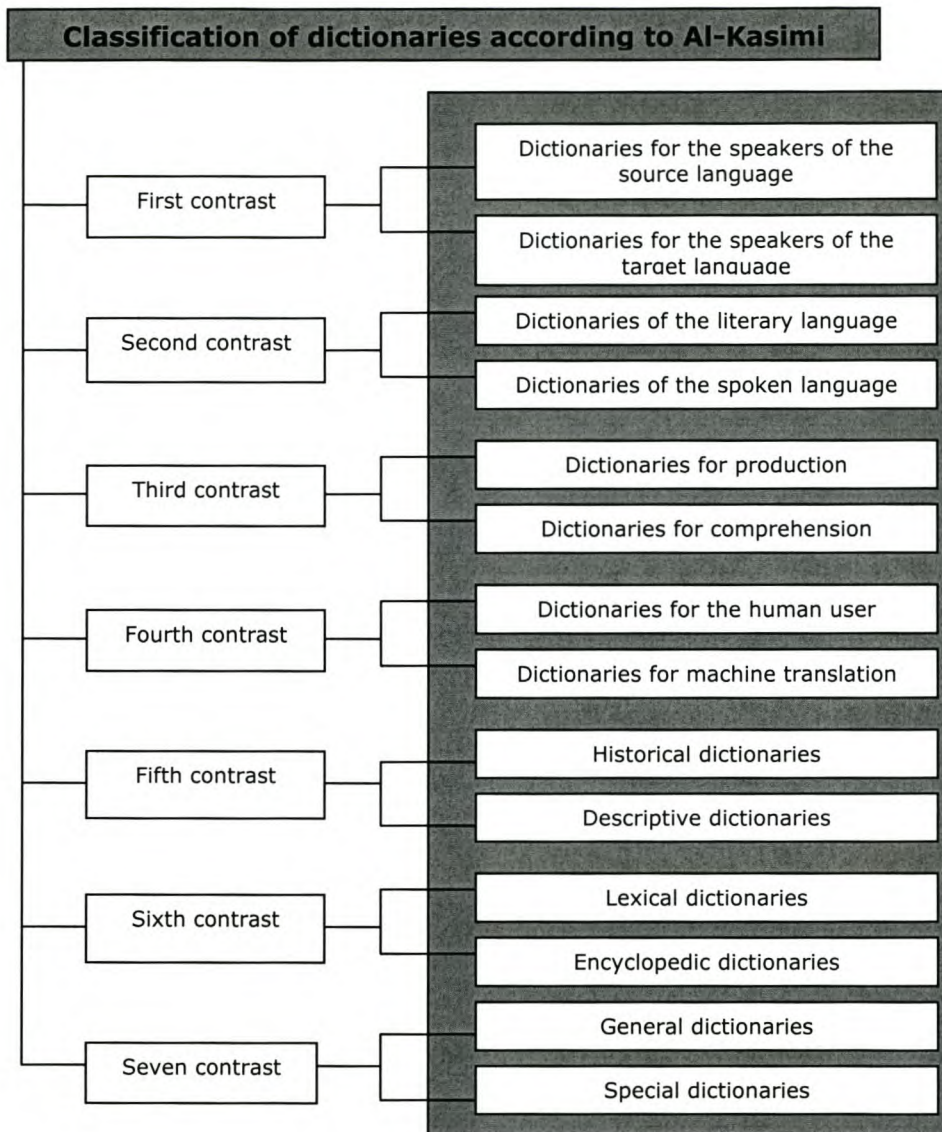
4. Dictionaries for the human user vs. dictionaries for machine translation. There are many essential differences between a dictionary intended for human users and one intended for machine translation. For instance, while the former is expected to provide only necessary information which the user needs, and which information is usually about the language that is foreign to the user, the dictionary designed for machine translation must contain much more detailed grammatical information about both languages.

5. Historical dictionaries vs. description dictionaries. The defining characteristics of historical dictionaries are: (a) the sources of the dictionary are written material or records that belong to the past of a language, and the information it presents on pronunciation is based on those records. As a result, the dictionary will include obsolete expressions. (b) When the dictionary provides etymologies, it leaves description and goes to history. (c) Quotations cited in the historical dictionary are limited to a specified past period or specified periods of the language. (d) The historical dictionary arranges the senses of its entries in such a way as to show how meanings have developed from one another.

6. Lexical dictionaries vs. encyclopaedic dictionaries. Both encyclopaedic dictionaries and encyclopaedias include encyclopaedic data. While in encyclopaedias the data is presented under fewer general topics, in encyclopaedic dictionaries it is distributed under a large number of related headings.

7. General dictionaries vs. special dictionaries. A general dictionary is one that attempts to cover the whole lexicon of the language whereas a special dictionary deals with one sector of the lexicon. In a general dictionary, the vocabulary of all fields of knowledge should be represented and the reading of the public for whom the dictionary is intended should also be sampled. On the other hand the purpose of the special dictionary is to help the user to acquaint himself with the meaning of the jargon and the terminology of a special field.

Al-Kasimi's classification can be schematised as follows.



1.2.6.5. Zgusta's classification

In his classification Zgusta distinguishes two main types of dictionaries: encyclopaedic dictionaries and linguistic dictionaries. According to him, encyclopaedic dictionaries are primarily concerned with the denotation of the lexical unit (words). They give more data about the extra-linguistic world, physical or non-physical, and they are arranged in the order of the words by which the segments of

this extra-linguistic world are referred to when spoken about it. Zgusta's classification does not focus on encyclopedic dictionaries. His main concern is related to linguistic dictionaries that he has divided into four main categories according to different criteria:

- The first division of linguistic dictionaries is made between diachronic dictionaries and synchronic dictionaries. Diachronic dictionaries are primarily concerned with the history, with the development of words, both in respect to form and meaning. There are two types of diachronic dictionaries: historical and etymological dictionaries. Historical dictionaries focus on the changes occurring both in form and meaning of words within the period of time for which there is historical evidence at hand, while etymological dictionaries focus their interest on the origin of the words. As far as synchronic dictionaries are concerned, their task is to deal with the lexical stock of a language at one specific stage of its development.
- The second division of linguistic dictionaries discerns general dictionaries from restricted or special ones. For restricted or special dictionaries the compiler decides to focus only on a certain part of the total lexicon of the language. The restriction can be based on any perceivable (or supposed or postulated) variation of the language, on any classification of its texts, or on any principle or combination of principles determined by the author of the dictionary. As far as general dictionaries are concerned, they are mainly dealing with the general language. Within the category of general dictionaries, it is necessary to distinguish two types: standard-descriptive dictionaries and overall-descriptive or informative dictionaries.

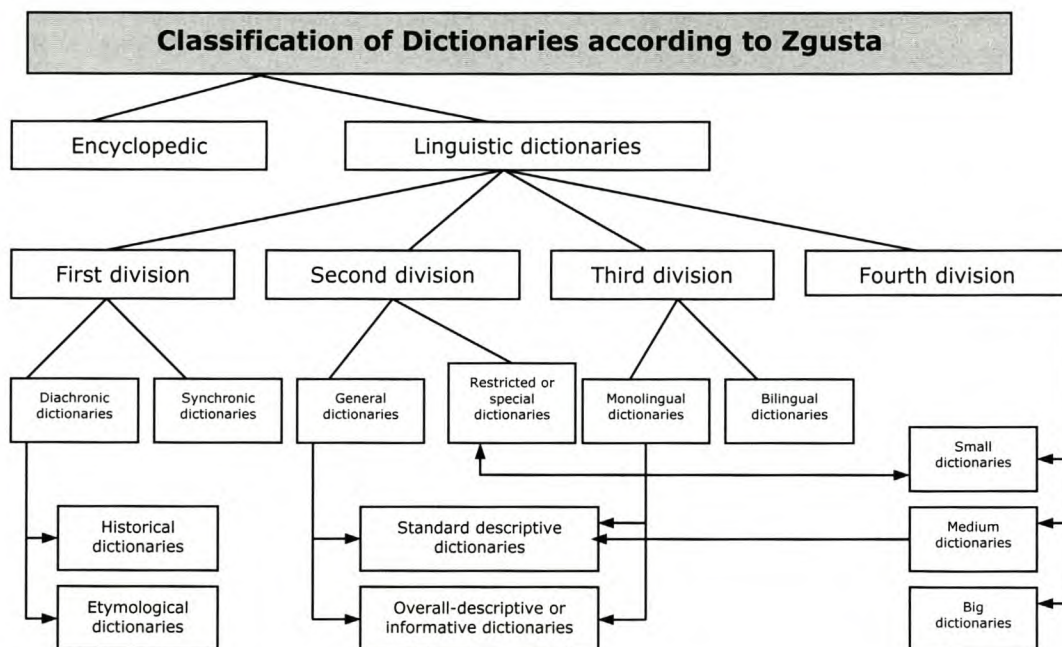
Standard-descriptive dictionaries can be characterized as descriptive dictionaries of the standard (national) language as it is used at the point in time when the dictionary is being compiled. Overall-descriptive dictionaries differ from standard-descriptive ones in the fact that they describe much more than the standard (national) language as it is used at the point of time of the compilation; they are not concerned with the future usage. They are primarily used by people who wish to find information about a word they do not understand when reading (or hearing) a text.

- The third division of linguistic dictionaries made by Zgusta concerns the number of languages in a dictionary. In fact, he makes the difference between monolingual dictionaries and bilingual ones. Where monolingual dictionaries deal with one, bilingual ones are dealing with two languages. It is important to retain, according to Zgusta, that a standard-descriptive, an overall-descriptive or an academic dictionary is usually monolingual.

- The fourth and last division made by Zgusta regards the size of dictionaries. Indeed, according to him dictionaries can be small, medium or big. Small dictionaries usually have no quotations or examples, so that its generative power is rather low. The vocabulary of the small dictionary is reduced in comparison with the medium one. The reduction usually eliminates not only the obsolete and similar lexical units, but also the less frequent and unimportant lexical units that belong to the standard national language and the less frequent senses of the polysemous lexical units. The medium

dictionary is usually of a rather more standard-descriptive character, because it saves space above all at the expense of the obsolete and dialectal lexical units and occasionally, less regular applications, so that it may be expected to list the lexical units of the standard (national) language in its contemporary form. As far as big dictionaries are concerned, they try to be as exhaustive as possible.

Zgusta's classification can be schematised as follow:



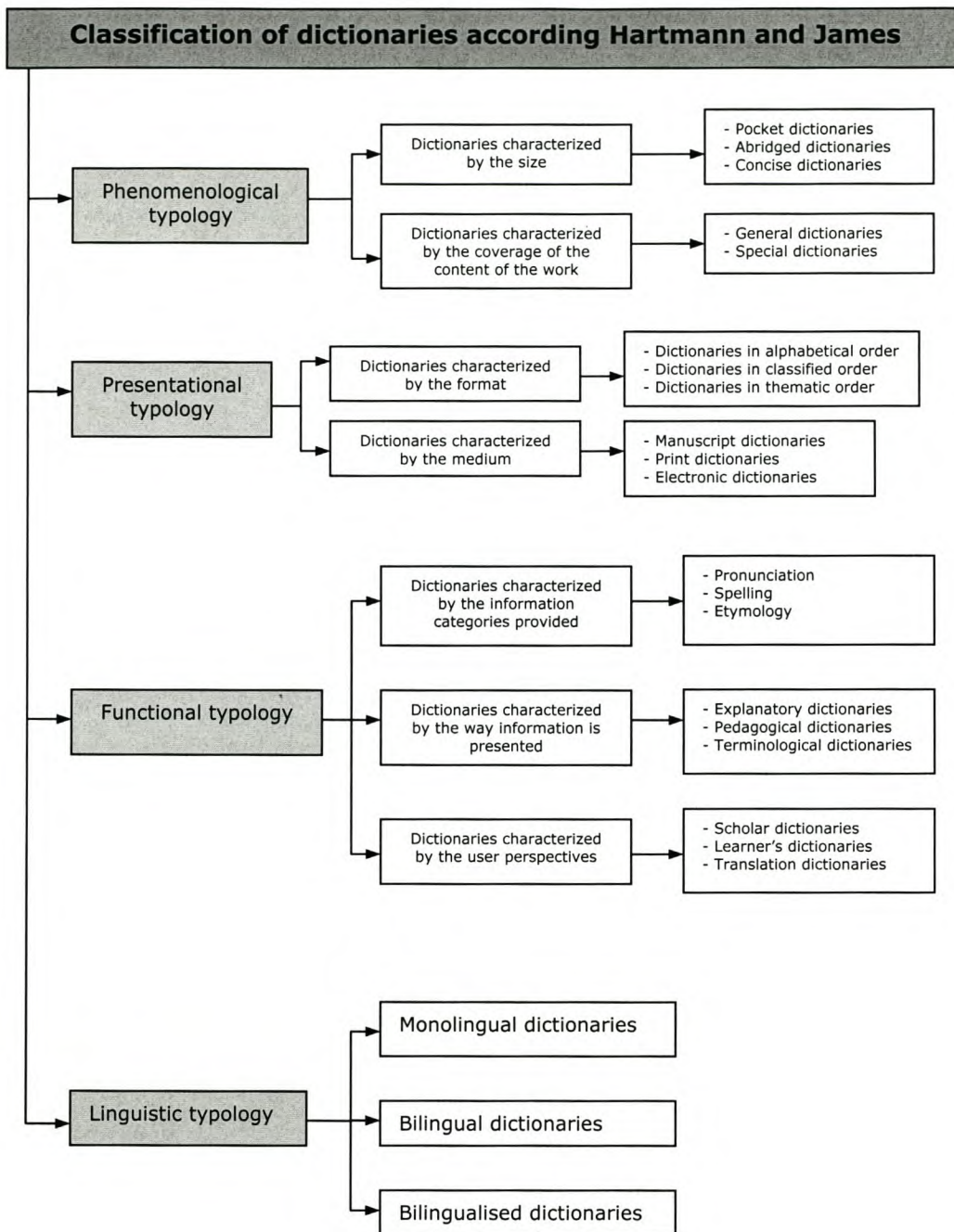
1.2.6.6. Hartmann and James' classification

In the introductory section of *Dictionary of Lexicography* (1998) Hartmann and James have made a general classification of dictionaries based on different aspects. According to Hartmann and James these different types of dictionaries have developed over a considerable period, in response to linguistic and cultural demands, and as a result of changes in the use and availability of

communicative media. They have classified dictionaries into four typologies or categories:

- The first category that they call the phenomenological typology is based on formal features, which take into account compositional characteristics such as size (e.g. pocket dictionaries, abridged dictionaries, concise dictionaries, etc.) or coverage of the content of the work (e.g. general dictionary, specialised dictionaries, etc.).
- The second category that they call the presentational or tectonic typology focuses on the format (e.g. alphabetical, classified, thematic, etc.) or medium (e.g. manuscript, print, electronic, etc.) of the dictionary.
- The third category that they call the functional typology, is based on the contextual uses of the dictionary, focuses on the information categories provided (e.g. pronunciation, spelling, etymological, etc.), and the way these are presented (e.g. explanatory, pedagogical, terminological, etc.) within the perspective of the target user (e.g. scholarly, learner's, translator's, etc.).
- Finally the fourth and last category, that they call the linguistic typology, focuses on the language(s) of the dictionary (e.g. monolingual, bilingual, bilingualised, etc.).

Hartmann and James' classification can be schematized as follows:



1.2.6.7. Béjoint’s classification

The last classification I am going to present is the one given by Henri Béjoint (2000) who maintains that the basic question for anyone who attempts a typology is whether to derive categories from the observation of existing dictionaries or to create categories

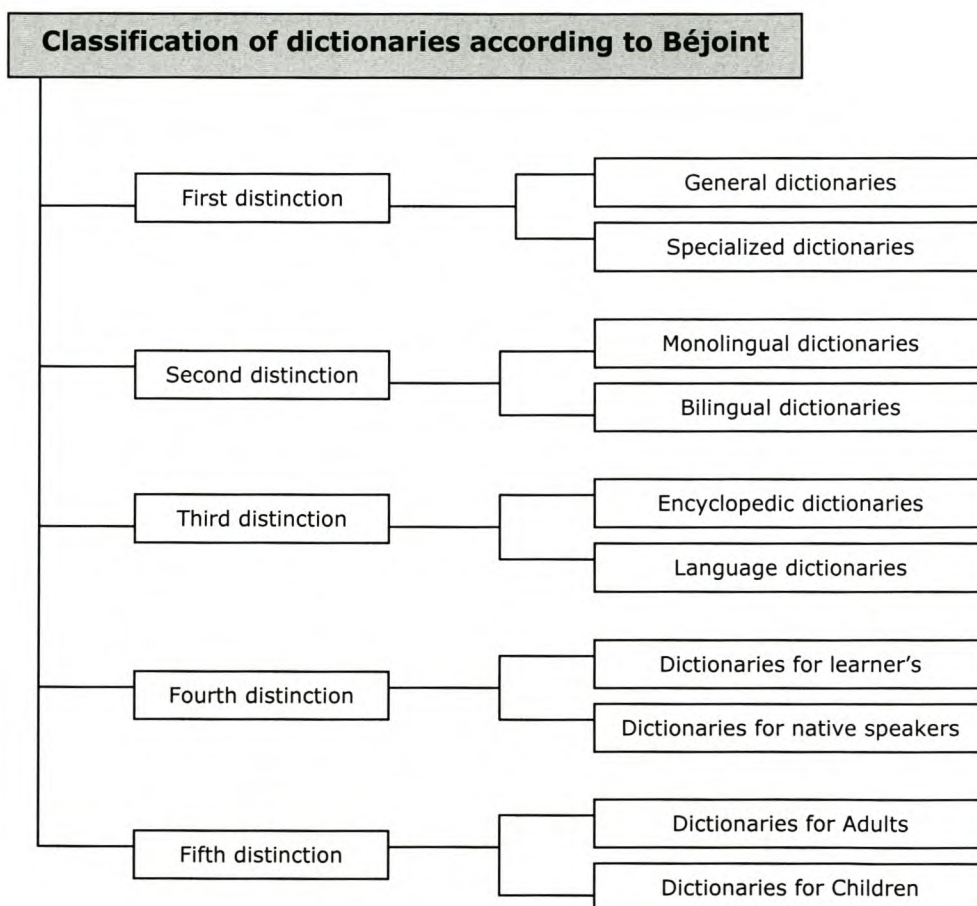
in theory and then see how existing dictionaries fit into them. Béjoint estimates that when dictionaries are categorized according to the existing types, one should call that a classification. However, when the categorisation of dictionaries is made according to a created theory it is called a typology.

Béjoint has classified dictionaries by making five broad distinctions between them:

- The first distinction is made between general and specialized dictionaries. A dictionary can be general or specialized in its macro- or microstructure or, in both. The macrostructure can be called general if it includes all, or a representative section, of the elements of a lexicon, even the obsolete and archaic, and also all its varieties in synchrony, and if it has all parts of speech as entry-words. A specialized macrostructure is one that is restricted to one variety, a dialect, technical jargon, slang, etc. or to one type of entry-word: verbs, adjectives, idioms, or proper names.
- The second distinction is made between monolingual and bilingual or multilingual dictionaries. Monolingual dictionaries are those in which the language of description is the same as the language being described. A bilingual dictionary uses two different languages: one as the object of description and one as instruments of description.
- The third distinction is made between encyclopedic and linguistic dictionaries.
- The fourth distinction is made between foreign learners' and native speakers' dictionaries.

- The last and fifth distinction is between dictionaries for adults and dictionaries for children. Dictionaries that are designed explicitly for children have a shorter word-list consisting of the most important words of the language.

Béjoint's classification can be schematized as follows.



1.2.6.8. Remarks

I share Béjoint's (2000:37) point of view that "all the dictionary typologies proposed by various authors over years show that it is impossible to classify dictionaries in a way that would be both orderly and applicable to all societies. Dictionaries come in more varieties than can ever be classified in a simple or unique

taxonomy". Therefore, there are no standard criteria whereby dictionaries can be classified. However, it is important to ensure that any classification must always somehow be related to the main characteristics of dictionaries and/or their users.

1.2.6.9. Classification of bilingual dictionaries: a new vision

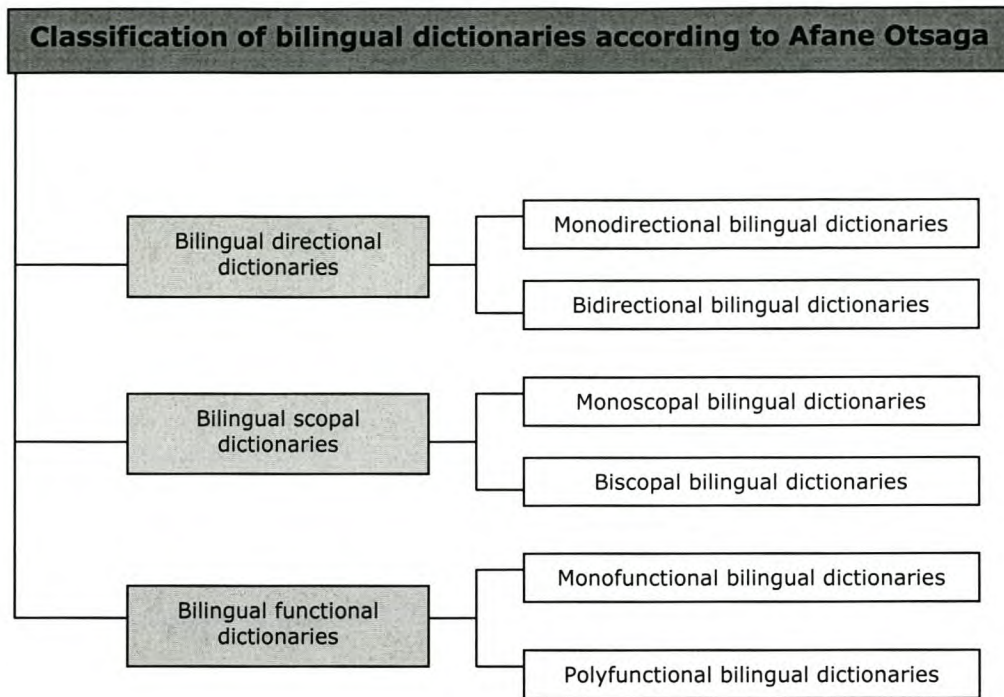
Most dictionary classifications up to the present often include all types of dictionaries (i.e. monolingual, bilingual, encyclopaedic, etc). Al-Kasimi has been one of the first researchers who proposed a dictionary typology based essentially on one type of dictionary, i.e. bilingual ones. As seen in that classification, dictionaries have been differentiated according to their contrasts. In accordance with the typological model I propose and by using Al-Kasimi's criterion, bilingual dictionaries (the main concern in this dissertation) can also be classified according to three other types of contrasts:

1. Monodirectional bilingual dictionaries vs. bidirectional bilingual dictionaries. Mono- and bidirectional here should be understood in Hausmann and Werner's way. Monodirectional bilingual dictionaries are targeting only one category of users, for instance those who use dictionaries as mother tongue speakers of the source language. Bidirectional dictionaries are targeting two categories of users, e.g. mother tongue speakers of the source language and mother tongue speakers of the target language.
2. Monoscopal bilingual dictionaries vs. biscopal bilingual dictionaries. Monoscopal bilingual dictionaries involve only one way in the relationship between the source language and the target language (e.g. from A to B). Only one section of the

dictionary is presented; e.g. lemmas in the source language are presented with their translation equivalents in the target language. Biscopal bilingual dictionaries include both ways, from the source language to the target language and from the target language to the source language (e.g. from A to B and from B to A). In clear terms, two sections of the dictionary are presented with lemmas and their equivalents given for both members of the language pair in two separate alphabetical components.

3. Monofunctional bilingual dictionaries vs. polyfunctional bilingual dictionaries. The function here particularly involves the purpose of the dictionary; the main reason why it is compiled. Monofunctional bilingual dictionaries will have one main general function, e.g. helping users to produce sentences in the target language. In that case, many examples and text illustrations will be included in the dictionary. Polyfunctional bilingual dictionaries will have at least two or more main purposes, e.g. helping users to get general knowledge in the target language by including general encyclopedic data, helping the standardization of one of the members of the language pair by using only one dialect in the dictionary, helping the user to improve the comprehension of the target language by including extensive grammatical data, etc.

The proposed classification can be schematized as follows:



1.2.6.10. The proposed dictionary model: A typological hybrid

I entirely agree with Gouws (2001:80) who said that new dictionary projects should not see themselves bound to the existing typological models but should have the freedom to create innovative typological models. According to him, Wiegand (1996) has given an example of a dictionary article containing elements typical of both a monolingual descriptive and a bilingual dictionary. Hartmann (1994) also works with a similar approach. Gouws (2001: 81) adds that the secondary comprehensive lexicographic process of each project could consider the possibility of starting their lexicographical endeavours with the compilation of a hybrid dictionary which could satisfy some of the needs for a monolingual descriptive and some of the needs for a bilingual dictionary. The model I propose should combine the following features:

-
- It will be a bilingual dictionary (i.e. two languages), because it will present French items with their translation equivalents in Fang. However, several features generally found in monolingual dictionaries will be added, e.g. a paraphrase of the meaning of lemmas, a description of the equivalents, etc. In the majority of bilingual dictionaries for instance, only the translation equivalents of lemmas are given. Nevertheless, sometimes the users do not know the real meaning of these lemmas. They usually need to go back to a monolingual dictionary to find out. The model I propose should save them such an exercise.
 - The dictionary scope will be monoscopal, as it will present only one direction of the relation between the language pair (From French to Fang). As Gouws said (2001:81) this hybrid dictionary will be monoscopal in the treatment presented in the central list because items from only one language will be lemmatised. Gouws estimates that an innovative approach could add some biscopal features to the dictionary and make the dictionary poly-accessible.
 - Its perspective will mainly be synchronic, because the selected lemmas will be coming from a database collected at a specific time or epoch. This dictionary will be aimed at the description of language at a given stage of its development without being concerned with historical changes that occur. However, some diachronic features may also be included if necessary. In fact, some aspects regarding the development of lexical items and their meaning may be included.

- The content will be both linguistic and extra-linguistic, because it will give all types of information and the target users should have access to a maximum of data concerning the language to be learned.
- Its purpose will be primarily prescriptive or normative, because it is a standard dictionary. That means it is a reference for the users and presents the norm in which they will be judged in their language use. Because it will be the first Fang standard dictionary, some descriptive and comprehensive features must also be included.
- It will have an alphabetical order, because it is the easiest way to use a dictionary and the way in which most users are familiar.
- It will be directed at both native and foreign speakers, because it will be used by young Fang speakers (as foreign speakers) and by any French speaking person (if French is the source language, anyone who knows French could easily use this dictionary like a native speaker). It will be a bidirectional dictionary directed at young Fang speakers and any French speaking person.

Within the typological category of bilingual dictionaries, the standard translation dictionary can be regarded as the best option to be employed as an instrument of standardization. Such a dictionary is desperately needed in Gabon. However, the compilation of such a dictionary has to be preceded by the formulation of a model, which includes the relevant dictionary plan, emphasising its standardisation function. This model has to reflect

the latest research in the field of metalexigraphy and must have a sound theoretical basis. The problem of language standardization, the role that the standard dictionary can play in the language standardization process and the formulation of a theoretically motivated model for a standard dictionary will be the focus of this research. This model will also identify different ways in which the dictionary could be used as an instrument in the standardization of Fang

In this research project a theoretical model will be devised for a standard translation dictionary with French and Fang as language pair. It will be designed for a monoscopal dictionary (from French toward Fang). The monoscopal orientation is motivated by the Gabonese context (the strong French presence) and by the fact that it should be easier for the target group (young Fang speakers) to retrieve information from a translation dictionary by using French as search route.

1.2.6.11. Standard features

Standard dictionaries can be regarded as products resulting from a well-established lexicographical environment. Standard dictionaries usually are single volume products in which a synchronic and normative approach prevails. The macrostructure represents the standard variety of the treated language although a number of high frequency usage items from non-standard varieties will also be included (Gouws, 2001:76).

Standard dictionaries primarily endeavor to reflect the standard variety of the target language(s). Therefore, they have to be normative in their approach. Standard dictionaries are the linguistic reference sources most commonly used by the average

member of a speech community and these users rely on the dictionary as an authoritative source of linguistic information (Gouws, 2000:15). Where a language does not have an established collection of dictionaries, the typical typological pattern will be a priority to compile one or more standard dictionaries. This type of dictionary plays an important role in the standardization process because users rely on the dictionary as an authoritative source (Gouws, 2000:15). Where a language is not yet fully standardized, it is customary to compile a standard translation dictionary before compiling a standard descriptive dictionary. In the Fang situation, even if some dictionaries exist in this language, they did not really influence the use of Fang in Gabon, because they were not and still are not available to the average member of the speech community. The dictionary model I propose will really be the first one, which should be used by the target group. To be a real standard model, the French/Fang bilingual dictionary should also include standard features in both languages. The selected lemmas must be presented in standard French, and the equivalents must primarily be given in the proposed standard dialect of Fang. The proposed model will also give some new approaches concerning the spelling, the phonetic alphabet and the phenomenon of tones that play an important role in the standardization process.

1.2.7. Dictionary functions

1.2.7.1. General remarks

In bilingual lexicography, it has become commonplace to contrast "active" and "passive" dictionaries or dictionary functions. Kroman, Riiber and Rosbach (1984:207) define an active dictionary as a monofunctional dictionary from the native language toward the

foreign language (LI > L2), and a passive dictionary as a monofunctional dictionary from the foreign language toward the native language (L2 > LI). Active dictionaries are used to translate from the native to the foreign language, while passive ones serve the opposite function, i.e. translation from the foreign to the native language. According to Tarp (2002) experience has shown that the determining element in a dictionary function is the user situation. There is, for example, much more difference between a dictionary conceived for text production in the native language and one conceived for translation into a foreign language than between a mother-tongue production dictionary conceived for adult users and one conceived for school children. For that reason, functions are frequently named after the corresponding user situations. Tarp specifies that the lexicographic functions can be subdivided into communication-orientated and knowledge-orientated in correspondence with the respective main types of user situations. The most important communication-orientated functions are as follows:

- To assist the users solving problems related to text reception in the native language.
- To assist the users solving problems related to text production of texts in the native language.
- To assist the users solving problems related to text reception in a foreign language.
- To assist the users solving problems related to text production in a foreign language.
- To assist the users solving problems related to translation of texts from the native language into a foreign language.

- To assist the users solving problems related to translation of texts from a foreign language into the native language.

The most important knowledge-orientated functions are:

- To provide general cultural and encyclopaedic information to the users.
- To provide special information about the subject field to the users.
- To provide information about the language to the users.

1.2.7.2. Functions of the proposed dictionary model

In the situation of young Fang speakers, the classification of French either as a native or a foreign language is not so easy. It is true that French is a partial foreign language in Gabon, but at the same time it could be considered as a native language for young Fang people. Indeed, for most of them, French is their mother tongue, because it is the first language they have started to speak, the language that they have acquired naturally as children, as opposed to one learned later, e.g. through formal education (Matthews, 1997:238).

Because there is a strong affiliation between French and the bigger part of the target group, the model I propose will be both an active and passive dictionary. It is active because a large part of the target group does have French as native language. The dictionary could also be used by anyone who has French as native language, and passive because French is a partial foreign language in Gabon and one part of the target group does not have French as native language. The passive function can also be justified by the fact that any Fang speaker who has some knowledge in French could also use the dictionary model.

As far as the communication-orientated functions are concerned, the proposed dictionary will:

- Assist the target-group to solve problems related to the reception of texts in Fang.
- Assist the target-group to solve problems related to the production of texts in Fang.
- Assist the target-group to solve problems related to the translation of texts from the French into Fang.

For knowledge-orientated functions, the proposed dictionary will:

- Provide general cultural and encyclopaedic information to the users.
- Provide special information about some specific subject fields, particularly those related to academic purposes.
- Provide information about Fang.

1.2.8. Survey of some aspects of some theoretical models and approaches

1.2.8.1. Introduction

In this research the model will focus on the structure, the contents and the presentation of data relevant to the goal. This model will use the theory of Hausmann and Wiegand, particularly those components of Hausmann and Wiegand's theory that are relevant to translation dictionaries. The model will also refer to a range of other works that have handled the same issues in lexicography, e.g. Landau (2001), Prinsloo (2000), Zgusta (1971), Kachru (1980), and Gouws (1996).

1.2.8.2. Landau and the material collection

Sidney Landau (2001) believes that the first stage in compiling a corpus is to decide why it is needed and what its purpose will be. He adds that the nature of the corpus will depend on the size and the scope of the dictionary for which it is designed, its intended audiences; and the resources (in both people and money) that can be marshalled for the compilation. Landau estimates that, for the sake of efficiency, a corpus should be no larger or more complex than it has to be to fulfil its functions. Landau suggests some ways in which the material should be collected, the domains of collection, the fields of collection and other different types of sources from where the data could be collected. Landau's approach, as far as the material collection is concerned, will form an important basis for the collection policy I intend to propose.

1.2.8.3. Hausmann and Wiegand and dictionary components

Hausmann and Wiegand (1989) proposed a theory for the component parts and structures of a general monolingual dictionary. According to Hausmann and Wiegand's theory, a monolingual dictionary must have the following major components:

- The **word list** constitutes the main part of the book (dictionary). It has several important units and structures. There may be a *central word list* and several additional word lists. The basic unit of the word list is the *treatment unit*, which results when a form mentioned and information relating to that form, are brought together. According to Hausmann and Wiegand, the way in which a form and related information are brought together is the *addressing procedure*. Each information *item* is addressed to a form called the *address*.

- Hausmann and Wiegand stress that the most important address is the *lemma*, also called *entryword* or *headword*. It belongs to the **access structure** of a dictionary. The *access structure* is the search route followed by a dictionary user to reach a *lemma-sign* or specific data in a dictionary. A distinction must be made between the *outer access structure* and the *inner access structure*. The outer access structure is the search route guiding the user toward the appropriate lemma-sign. The inner access structure is the search route within the article guiding the user toward specific data or a specific entry.
- All the *lemmata* (or lemmas) constitute the **macrostructure** of a dictionary. If the dictionary has one word list only, the *outer access structure* and the macrostructure are identical. If the dictionary has additional word lists, the central word list and the additional word lists are all partial structures of the macrostructure.
- The lemma and the whole set of data items which are addressed at the lemma, form the *dictionary article*. The data structure within the article is called the **microstructure**.
- The **mediostructure** is the system of reference employed in a dictionary that leads a user from one entry to another. A lexicographer refers the dictionary user from a reference position to a reference address.

Even if the components identified by Hausmann and Wiegand refer to a monolingual dictionary; I will use and adapt them for my proposed model, particularly aspects relevant to the macro-and

microstructure, the access structure, the addressing structure and the mediostructure.

1.2.8.4. Gouws and the outer texts of the dictionary

According to Gouws (2000:21), one of the focal points in the most recent lexicographic discussions concerns the structure of dictionaries. The motivation for this concern is the effort to create quicker and easier access to dictionaries and to ensure that users can find and retrieve the information they are looking for. He adds that successful dictionary use relies on a dictionary text, which allows easy access to the required data. To be able to utilize such a text to the full, it is necessary to be familiar with the editorial system of the dictionary. Gouws clarifies that besides the central list, which is the section of the dictionary that includes the selection of lexical items and their lexicographic treatment, a dictionary should have another component, which has to contain an explanation of the editorial system. This is usually presented in a component that precedes the central list. This component may also contain other data concerning grammatical information, notes on pronunciation, etc. Many dictionaries also include some data in addenda, which follow the central list. These addenda can contain lists of abbreviations, symbols, units of measure, geographical names, etc. He calls the addenda, which precede the central list, the **front matter** and the addenda, which follow the central list, the **back matter**. Front and back matter texts form the outer text and help constitute the frame structure of a dictionary. Gouws adds that the use of outer texts allows the lexicographer to include entries, which would not typically appear in a linguistic dictionary. In the back matter section the outer texts could include various lists, e.g.

the names of countries, languages, symbols, abbreviations, etc. Gouws' approach to the outer texts in the dictionary will give me the opportunity to propose my vision about the contents and the role of the front and the back matter in a French/Fang translation dictionary.

1.2.8.5. Zgusta and language standardization

Zgusta (1971: 174) notes that a "difficult situation is found in the cases where there is no standard national language, or no written forms or no important literature, but where there are only spoken forms of several dialects. There are different developments possible in such a situation". Zgusta gives two examples where languages have been standardised. The first one was done in a large geographical area with several cultural and political centres, while the other was carried out in a small geographical area with a single cultural or political centre. According to him, in the majority of cases, it happens that during the development of a dialect into a national language some features of other dialects find their way into the standard national language. He adds that, developments have taken place in many languages without any political or economic preponderance and such developments are still possible today. The examples quoted by Zgusta have some aspects identical to the Fang situation. Zgusta's experience will help me find an adapted route in the Fang standardization process.

1.2.8.6. Gallardo and the role of the dictionary in the language standardization process

Gallardo (1980: 62) thinks that "the structural properties of the standard language are the areas where the influence of the dictionary is most visible". According to him, "the dictionary is a

crucial element in the way the frame-of-reference function of the standard language manifests itself, and it is intimately related to the awareness-of-the-norm attitude among the members of the speech community". Gallardo also proposes some ways in which dictionaries could play an important role in the language standardization process. I will apply those proposals to the Fang situation and will show how the French/Fang dictionary I am proposing could help to standardize Fang.

Apart from the theoretical models of Landau, Hausmann and Wiegand, Gouws, Zgusta and Gallardo, I will also use other available works on material collection and analysis, translation dictionaries and/or standardization, e.g. Bo Svensén (1993), Prinsloo (2000), Rey-Debove (1971), Dubois and Dubois (1971), Béjoint (2000), Kennedy (1999), etc.

CHAPTER 2: Fang and its dialects

What one calls "the Fang language" can be considered as a whole of speech forms very close, but a bit different and which belong to a vaster linguistic group that does not have an autochthon denomination. It is the group Pierre Alexandre calls "Beti-Bulu-Fang" and that, with some other authors, he also designates by the term "Pahouin" (Ondo Mbiame, 1992; translated from French).

2.1. Preliminaries

The term Fang designates the language as well as the speakers of this language (plural Befang). Fang is a Bantu language, which belongs to the A70 group also called Yaoundé-Fang group and is designated by the acronym A75, according to the classification of Bantu languages made by M. Guthrie (1948, 1953 and 1967-1970).

Table 3: Group Yaoundé-Fang (A70)

| Acronyms | Languages | Locations |
|----------|---|---|
| A71 | Eton | Cameroon |
| A72 | Ewondo Mwëlë Bakja or Badja Yangafek | Cameroon Cameroon Cameroon Cameroon |
| A73 | Bebele Gbigbil | Cameroon Cameroon |
| A74 | Bulu | Cameroon |
| A75 | Fang | Gabon, Cameroon, Congo, Equatorial Guinea. |

In this classification Guthrie takes into account the typological and the geographical criteria of Bantu languages that are divided in two branches:

- The West branch or Western Bantu designates Bantu languages spoken in Cameroon, in Gabon, in the Republic of the Congo, in the Western half of Zaïre (actually the Democratic Republic of the Congo or DRC), in Angola and in Namibia.
- The East branch or Eastern Bantu designates Bantu languages spoken in the Eastern half of the DRC and in all countries from the East and the South of Africa (Alexandre, 1948).

The migration of the Fang population is dated around the 18th and 19th century. The source regions of Dja, Ivindo and Ntem are considered as the last step before the dispersion of the Fang population (Mouguiama-Daouda, 1995). They reached the North and the West coast of Gabon during the middle of the 19th century. The Betsi group (Fang-Atsi) would be the first one to attain the Gabonese territory in the Southwest by Mitzic, Medouneu or by the Okano, Abanga and Bokwe Rivers. They also reached Ndjolé and Lambaréné. Another group headed by the Nzaman (Fang-Nzaman) entered Gabon through the Okano and the Mvoug Rivers. They occupied the right part of Ogooué River and established themselves at Ndjolé and Booué. Two groups, the Ntumu group (Fang-Ntumu) and the Mvaï group (Fang-Mvaï) stayed around the Woleu and the Ntem Rivers. Some groups from the Fang migration remained in Cameroon and are today going under the names of Ewondo, Eton, Betsi and Bulu.

Ella Ella (2000:14-15) presents two other versions as regards the migration of Fang people:

-
- The first version, more abstract, is symbolized by the so-called "hole of Adzap" legend. According to this legend, Fang people, leaded by a chief, came from an unknown place named "Odzamboga"¹. During this long journey they got stuck by a huge and long tree that they could not cross to continue their journey. This tree was called "adzap" and the only way to cross it was to make a hole through it. They did it after tenacious and tiring efforts. After this hard test they divided themselves in several groups who took different ways and reached the current locations of Fang people all over Africa. This version gives the mythical origin of Fang people.
 - The second version situates the origin of Fang migration toward the Gabonese coasts throughout the Ogooué river around the beginning of the 18th century and their implantation in Gabon can already been evaluated over more than 8 or 9 generations. According to Trezenem, as cited by Balandier, Fang people under the pressure of conquerors (Foulbé people) from Cameroon reached the Moyen-Ogooué (mid-Ogooué) at the beginning of the 19th century. Their arrival in the Estuaire must be situated around 1860. According to Duchailu, also cited by Balandier, in 1856 there was already talking about the presence of a race of active and enterprising people who was trying to take possession of the littoral.

¹ That place has not been identified until now

Today Fang people occupy the northern part of Gabon. Their border is the Ogooué river line. They are separated at the frontier of Cameroon and Equatorial Guinea by a whole range of related population group (Ella Ella, 2000). A lot of work about Fang and its dialects has been done. The objective of this chapter is to present some of those works by trying to classify them in each dialect. The accent will, in particular, be on lexicographical and linguistic works that could play an important role in the compilation of my corpus. Religious missionaries and colonial administrators compiled the earliest works in Fang. Most of these works were done in Fang dialects located where these religious and colonial administrators settled in Gabon for the first time (i.e. Estuaire and Moyen-Ogooué). Since the creation of the Omar Bongo University in 1971 and the Department of Language Sciences in 1995, most works on Fang are produced by students and lecturers. The Raponda Walker² Foundation has also published some works in Fang.

2.2. Language and dialect

Despite the fact that the dichotomy language/dialect is not always easy to understand, those two notions have clearly been scientifically defined in the linguistic field. When dealing with the standardization of a language with dialects as Fang, it is important to make a clear dissimilarity of both notions in order to make a better understanding of the standardization process. If we observe any language carefully, we shall see that no two people, even if they are speaking the same

² The Raponda-Walker Foundation is an organisation that tries to publish all unpublished works left by the Gabonese researcher A. Raponda Walker. This Foundation is also working on the promotion of Gabonese languages.

language, speak it in an absolutely identical way. There may be only trifling differences, but they are observable. They concern pronunciation, preferences in sentence construction, knowledge of lexical units, their use, etc. In this sense, the linguistic system of a single speaker of a language is called his idiolect (Zgusta, 1971:164). According to Zgusta (1971:168), if we observe the idiolect of speakers of the same language who have lived together since their childhood in a closed geographical area, we usually perceive (1) a basis identity of the idiolects with the idiolects of other speakers of the same language from other areas (2) some isolated features which are to be found in single idiolects only and which render them individual and (3) some features which alone or in specific groups are neither individual nor observable in the whole language, but which are typical just for that area, and are the cause that it is, within the language in question, distinguished from other similar areas. The form of language spoken in such an area is then called a dialect.

Hartmann & James (1998:80) characterize the language as *the most basic means of communication between human beings*, when they identify the dialect as *a variety of a language*. Anthony Lodge (1993:15) considers that language variation derives not only from the different uses to which a language is put, but also from its users, i.e. from speaker variables such as their geographical origin, their social class, their ethnic group, their age, sex and so on. Anthony Lodge adds that the relationship between language and dialect is inclusive rather than exclusive. Language is a generic or superordinate term

embracing any number of dialects and styles. He illustrates his words with the following figure.

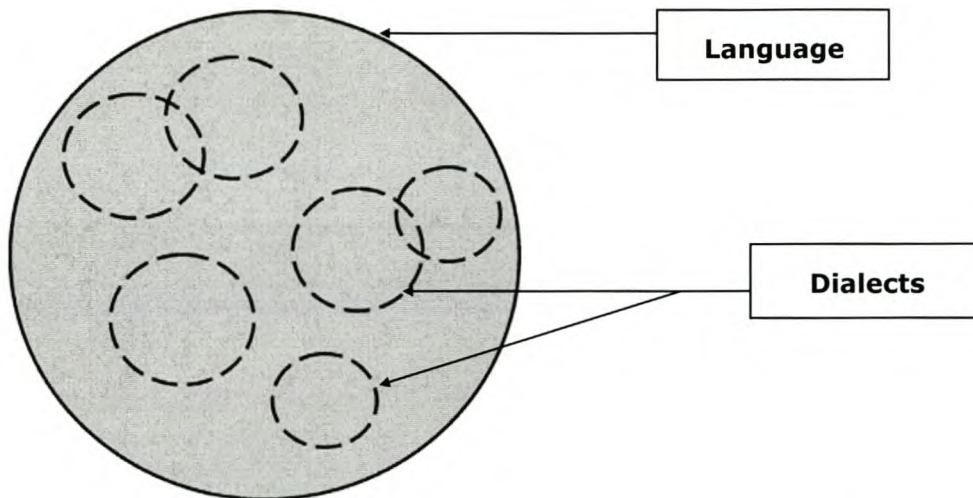


Figure 1: Relationship between language and dialects inspired from Anthony Lodge (1993).

The criterion for distinguishing dialect from languages is taken, in principle, to be that of mutual intelligibility³ (Matthews 1997:96). For instance, speakers of Afrikaans cannot understand Fang unless they have learned it, and vice versa. Therefore Afrikaans and Fang are different languages. But a speaker from Minvoul can understand one from Bitam⁴; therefore they speak different dialects of the same

³ The term intelligibility between dialects must be taken with caution. That is a matter of degree and ordinary usage. For instance Italian dialects (dialetti) are so called though many from the north and south are not mutually intelligible. By contrast Danish and Norwegian are called languages though speakers understand each other reasonably well (Matthews, 1997:97). The same situation can be observed in Gabon between some languages. For instance Yilumbu and Civili are identified as two different languages, but at the same there is a very high mutual intelligibility between speakers from both languages.

⁴ Minvoul and Bitam are two different regions of Gabon where Fang is spoken.

language (i.e. Fang). Petyt (1980:11) confirms that when saying, "*Using a language thus necessarily involves using one of its dialects*". Despite their mutual intelligibility, dialects from the same language remain differences (e.g. pronunciation, grammar, syntax and vocabulary) that allow speakers to distinguish them. Many reasons can explain the birth of these differences, although two of them are frequently identified: the need for a specific group to identify or distinguish itself from the whole speech community and the natural development of any language which evolves in different ways in different areas.

2.3. Fang dialects in Gabon

2.3.1. Introduction

Apart from the fact that Fang is widely spread into Gabon (actually in 4 provinces and 17 regions), it is also spoken in neighbouring countries (i.e. Cameroon, Equatorial Guinea, Republic of the Congo and Sao Tome). This means that varieties of Fang⁵ spoken in these countries are also counted as dialects of that language. However, in this work I will primarily be focusing on Fang dialects spoken in Gabon.

2.3.2. The Fang-Atsi dialect

Fang-Atsi is spoken by the Betsi⁶ population in 3 regions of Gabon: Lambaréné, Ndjolé and Bifoun. Fang-Atsi lexicography starts with the publication of a monoscopal bilingual dictionary *Dictionnaire fang-français* by R.P. Marling (1872). This work is followed some years

⁵ For instance Bulu is a Fang dialect spoken in Cameroon and its mutual intelligibility with Fang-Mvaï is recognised by the speakers.

⁶ Endogenous designation from Fang-Atsi speakers.

later by another monoscopal bilingual dictionary *Dictionnaire fang-français* by R.P. Lejeune (1892), which consists of 347 pages divided into two parts. The first part contains an overview of Fang grammar; the second part is the dictionary itself, which presents different lemmas in French followed by the translation equivalents in Fang.

One of the most important works in Fang-Atsi was compiled by the colonial administrator V. Largeau (1901). It is the *Encyclopédie pahouine*, which has about 4996 articles covering 699 pages (Nyangone Assam & Mavoungou, 2000). This book is the only encyclopaedia ever made in a Gabonese language. According to the author, the objective of this work was to give all colonial administrators exercising in Fang areas the opportunity to communicate directly with the Fang population without an interpreter⁷. The book is presented in three main sections.

The first section is a cultural (ritual, value system, mythology, origin, etc.) overview of Fang people, especially Fang-Atsi speakers in Lambaréné, Ndjolé, Bifoun and even Libreville and Kango. The second part contains the user's guidelines and the mini-grammar of the encyclopaedia. Some indications are made about the pronunciation system, the orthography system and the punctuation system, as well as some morphological data such as the parts of speech, word formation and the conjugation system of Fang. The third and last part is the encyclopaedia itself, which presents lemmas in French, in alphabetical order, followed by a description of the meaning of the

⁷ According to the history, the interpreters sometimes gave very bad translations that caused many problems between administrators and the local population.

lemmas, also in French, and then, the translation equivalents of lemmas in Fang, sometimes illustrated with examples. The Pastor Samuel Galley has compiled the most important work in Fang. It is a biscopal bilingual dictionary entitled *Dictionnaire fang-français et français-fang* (1964). This work, edited by Henri Messeiller and published after the death of Galley, consists of about 13 925 articles covering 588 pages (Nyangone Assam & Mavoungou, 2000). It is the result of the input from the Société des Missions Évangéliques de Paris in Gabon through its station of Talagouga in the region of Ndjolé.

Apart from these dictionaries, some works were also written in Fang-Atsi, particularly in the religious field. It is for instance the case with *Nten fan osua*⁸: *Premier livre de lecture en Fan* (1912), *Yesu: essai d'hamonisation des 4 Evangiles* (1908) and *Nten bya bi fan*⁹ = *Cantiques en langue fang* (1910), produced by the Protestant Mission of French Congo, French Protestant Mission and St François Xavier Mission respectively. The latest research in Fang-Atsi has been made by J. Mba Nkoghe (1979). It is a linguistic work presenting some phonological and morphological aspects of Fang. The data used for the production of this work were compiled in the Fang-Atsi areas (Ndjolé, Lambaréné).

The majority of lexicographical works in Fang has been made in the Fang-Atsi dialect. But these works are unknown to the greater part of Fang people, even Fang-Atsi speakers. This is either, because they

⁸ Qualified by the authors as the first reading book in Fang.

⁹ Hymns book in Fang.

are only used in specific fields (religious texts), or because, they are not available to a large public. Dictionaries for instance, are more available in European libraries (particularly in France) than in Gabon. That is why, despite all the important works published in Fang-Atsi, this dialect did not really influence the use of Fang in the rest of the country.

2.3.3. The Fang Mekè dialect

Fang-Mekè speakers can be located at the other side of the sea, in Metek ma vii chinchoua zone (on the West side of the Estuaire river) and in the area of Lalala (on the East side of the Estuaire river). According to her Fang-Mekè speakers are also called *mekè me Nkoma*¹⁰ because they are the group of Fang people that decided to leave and cross the Komo river when they reached the Estuaire. The name Mekè came from the Fang term "meke" (or mekè) which means "departure". This term today designates all Fang speakers in the Estuaire area, even those ones who did not cross the Komo river.

Fang-Mekè is recognised as the Fang dialect in which most religious works have been published in Fang. Indeed, the Bible and some of its parts (particularly the Gospels) were translated in Fang-Mekè (particularly by the Catholic religion). This situation can be explained by the fact that Fang-Mekè is spoken in the province of Estuaire (Libreville, Kango, Ntoun and Foulenzem), which was the first Gabonese region to be in contact with European missionaries and colonists. As far as the Catholic Church is concerned, because

¹⁰ Those ones who went to Nkoma or Komo River.

Libreville is the siege of the Clergy, most religious texts translated in Fang were done in the dialect spoken in this area.

The most famous work known in Fang-Mekè is the translation of the Bible by the first Gabonese Catholic Archbishop, François Ndong (1962), under the title of *Ntem Wam*. Until today, this book is still used in Fang Catholic Churches all over Gabon. The Archbishop F. Ndong, who was a Fang-Mekè speaker, also influenced the compilation of various works in this dialect.

Some important religious works have also been made in the translation of the Gospel in Fang-Mekè. In fact, *Évangile de Jean: traduction dans la langue des Fan* (1910), *Évangile de Luc: traduction dans la langue des Fan* (1910) and *Évangile de Matthieu: traduction dans la langue des Fan* (1902) are the Fang translations of the Gospels of John, Luke and Matthew respectively.

In the same order *Catéchisme Fan* (1932), *Récits de l'Ancien et Nouveau Testament = Nten nzamoe* (1936) and *Syllabaire Fan* (1925) were compiled by the Catholic Mission of Libreville in Fang-Mekè, for the teaching of the Gospel in Fang. It is also with the help of the Catholic Mission of Libreville, that H. Trilles produced several works in Fang-Mekè, e.g. *Ntèn wa tar'éyé gele fang'ne fala étén nélang = exercices de lecture et d'écriture en pahouin et en français* (1898), *Livre de prières = NTEN WAM minkobe ma nè ge kobe né nzame* (1898), *katesism Nten wa vegele Fang Nsong Katolik = Catéchisme de la Doctrine Catholique* (n.d.). All these works also concern the teaching and reading of the Gospel in Fang.

The latest work in Fang-Mekè has been compiled by Nzang-Obame (2001) in *Esquisse d'un dictionnaire fang-Mekè*. In this research project she presents a theoretical model of a dictionary outline in Fang-Mekè.

According to the researcher the aims of this work is to compile a sample of a Fang-Mekè dictionary with all the features of a complete dictionary. All those illustrations show that numerous works (particularly religious ones) have been made in Fang-Mekè. But, the influence of Fang-Mekè in the use of Fang is only perceptible throughout the religious field, specifically the Catholic one¹¹. This influence does not go beyond that. One reason for this situation could be the fact that Fang-Mekè does not have many speakers, compared to the other Fang dialects.

In informal discussions, speakers from other Fang dialects are always saying that Fang-Mekè has so many external influences because of its early cohabitation with some other languages and cultures in the Estuaire¹² region. They think that Fang-Mekè is no longer pure Fang. My feeling is that, because of this "sentiment of impurity" concerning Mekè-Fang, the other Fang speakers are, unconsciously, rejecting this dialect. This could explain why Fang-Mekè did not impose itself beyond the religious fields.

¹¹ The Protestant religion in Gabon is under the influence of the Bulu, one of the Fang varieties in Cameroon.

¹² In the Estuaire region, Fang-Mekè cohabits principally with Myénè and French, but also with languages from Gabonese and non-Gabonese people living in the capital (Libreville) and towns around.

2.3.4. The Fang-Mvaï dialect

Fang-Mvaï is the sole Fang dialect spoken in only one region of Gabon. Indeed, this dialect is only spoken in Minvoul or Haut Ntem region. I inventoried three works produced in this dialect.

The first one is a linguistic dissertation made by N. Mba Nzué (1981), entitled *Esquisse phonologique du mvaï (parler de Minvoul)*. In this work, the author describes the phonemes that he has identified in Fang-Mvaï and the contexts in which these phonemes combine. The second and third works in Fang-Mvaï produced by J.E. Mbot (1970 & 1975), are the beginning and the continuation of the same research. In *Ebughi bifia, un mode de connaissance de la langue fang*, the author tries to show how the use of Fang in some specific ways could play a role in the acquisition of knowledge in Fang. In *Ebughi bifia, "démonter les expressions". Enonciations et situations sociales chez les fang du Gabon*, J. E. Mbot makes a parallel between the use of the language and the social status of the user.

The region of Minvoul has served as a main field of investigation for these works. Some people see Fang-Mvaï as the less mixed dialect of Fang and some of them see Minvoul as the region where the real¹³ Fang is still spoken. But none of these allegations has been proved scientifically. The small number of its speakers and the fact that only a few works have been made in Fang-Mvaï did not give this dialect the opportunity to influence the use of Fang beyond Minvoul.

¹³ They qualified this dialect as Fang-Fang, which means Fang from original Fang.

2.3.5. The Fang-Ntumu dialect

Fang-Ntumu speakers are the most important community of Fang in Gabon. This dialect, located in Oyem and Bitam, covers the largest Fang area in Gabon. Large amounts of work compiled in Fang-Ntumu are linguistic studies. Ndong Menini (1980) produced the first phonological analysis in Fang-Ntumu, entitled *Esquisse phonologique du ntumu*¹⁴. Some years later during the seminar of experts about the Scientific Alphabet of Gabonese Languages, Voltz (1989) presented an important study he made in Fang-Ntumu. This work included some historical and cultural aspects of Fang, as well as classification, location and particularly phonetic, phonological and morphological aspects of Fang-Ntumu.

The most important linguistic work ever compiled in Fang-Ntumu has been done by Ondo Mebiame (1992) in his doctoral dissertation, *De la phonologie à la morphologie du fang-Ntumu parlé à Aboumezok*¹⁵ (*langue Bantu A.78*). This work in 2 volumes, gives more details about the functioning of Fang in general, and Fang-Ntumu in particular. Phonemes, syllables, morphemes and supra phenomenal elements of Fang are treated with an unprecedented precision in this work. The work by Voltz and Ondo Mebiame will be important parts of my database. Some interesting linguistic works related to Fang-Ntumu and released by students have been done or are currently in progress at Omar Bongo University. It includes the case of the following works:

¹⁴ Phonological outline of Ntumu.

¹⁵ From phonology to morphology of Fang-Ntumu spoken in Aboumezok.

-
- *Etude contrastive des pronoms relatifs en Français et en Ntumu* (Ella Ella, 2000). This research project which is a contrastive study of relative pronouns in French and in Fang (Ntumu) tries to demonstrate the existence or not of relative pronouns in Fang-Ntumu according to some criteria like the structure, the morphology or the syntax of relative pronouns in both languages (French and Fang). This work also tries to show the difficulties that young speakers can encounter in a situation of bilingualism.
 - *La présence des faits suprasegmentaux du fang ntumu dans l'utilisation du français pour les enfants âgés de 10 à 12 ans* (Mekui Missang, 1998). In this work the researcher identifies the apparition of supra phenomenal features from Fang-Ntumu in the use of French, as well as the cause of these interferences. Young speakers from 10 to 12 years old that have Fang as first language and French as second language have been used as a test group in this study.
 - *Etude des compétences narratives chez les enfants bilingues Fang-Ntumu-Français d'Oyem (Psycholinguistique): Etude transversale sur le développement de la temporalité et de la connectivité chez des sujets âgés de 7 à 12 ans* (Bibang Meye, 2001). In this research, which has a psycholinguistic approach, the author tries to find out how tools that allow the production of coherent narrations in Fang-Ntumu and in French are developing among bilingual kids and how the

development of both linguistic systems manifest themselves among these children. Young Fang-Ntumu speakers from 7 to 12 years old were the target group of this work.

- In *Etude linguistique d'une oeuvre littéraire: Un mvet de Zué Nguema, chant épique Fang recueilli par herbert Pepper*, Maillard Ella (2000) studies Fang through a literary work. In this research the author shows how it is possible to seize the dynamic of a language by using linguistic tools presented in a literary work. The literary work in question was an epic story of Fang-Ntumu.

Beside linguistic works, Fang-Ntumu also has some works in other fields, e.g. *Le mvet, genre majeur de la littérature orale des population pahouines (bulu, fang, ntumu)* by G. & F. Togo Atangana (1995), which is about the oral literature in Fang culture. This work is written in French. This is also the case with *Sagesse et Initiation à travers les contes, mythes et legends fang* from Mvé Ondo (1991). In the audiovisual media, Fang-Ntumu is the most used dialect in the presentation of news in Gabonese national media (radio and television). Even some religious films, like *Jésus de Nazareth*¹⁶, are translated into Fang-Ntumu.

In economic, educational and social plans, Fang-Ntumu has occupied and still occupies an important place in the activities of Fang people in the Woleu-Ntem province. More than half of the Fang people in Gabon are from this province. Woleu-Ntem is the only

¹⁶ Jesus of Nazareth.

monolingual¹⁷ province in Gabon, and has 5 main regions (Minvoul, Bitam, Oyem, Mitzic, Medouneu). Because of this homogeneity, all interactivities between people are done in Fang. A lot of people from the rest of the province are leaving their original regions, going to Oyem and Bitam to study, solve administrative problems, do shopping, do business, work, etc. Those two regions offer more facilities as far as these activities are concerned in this province. Because of this situation Fang-Ntumu influences the use of Fang in the whole province. The most visible illustration in this regard comes from some families where the parents are speaking Fang-Mvaï for example, but their children are speaking Fang-Ntumu, because they were born in Ntumu areas, or they were still young when the parents moved in those regions.

2.3.6. Fang-Okak dialect

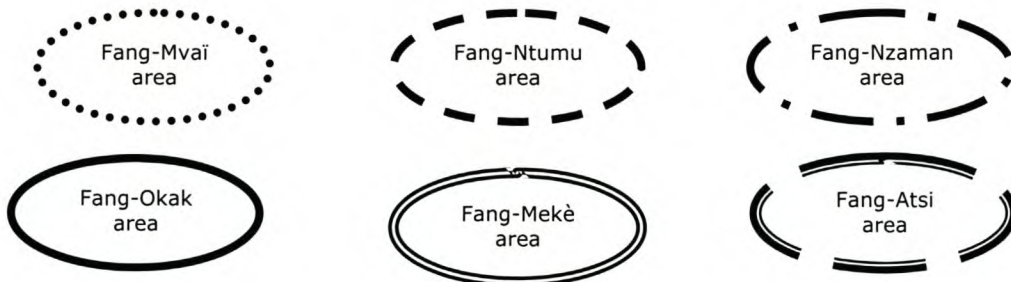
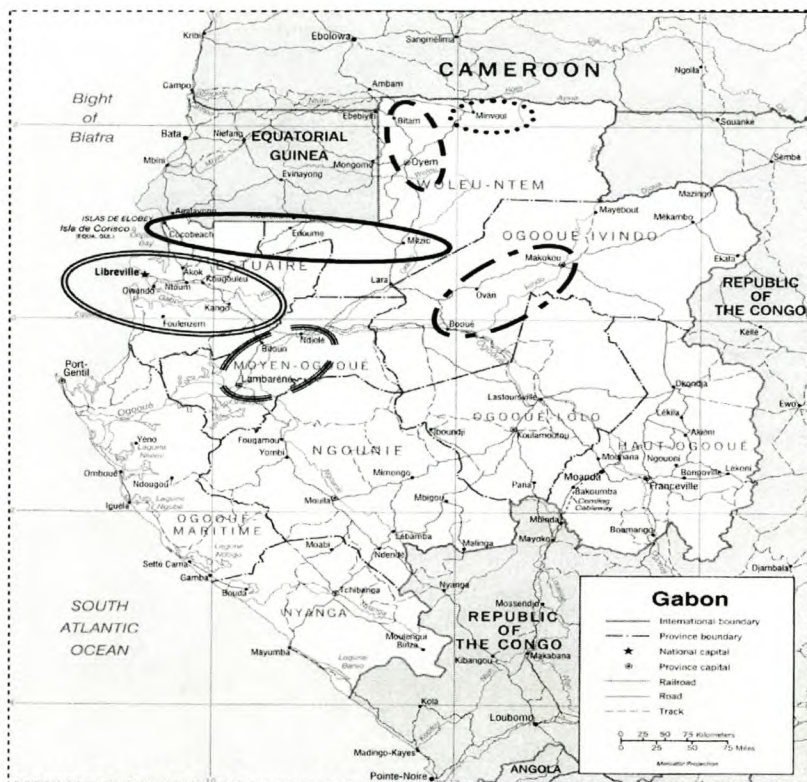
Located principally in the Mitzic, Medouneu and Cocobeach regions, the Fang-Okak is, maybe¹⁸, the sole dialect of Fang in which no linguistic or lexicographical work has been done. I think there are two reasons for that situation. The first one is the small number of speakers, and the second one is its geographical situation. In fact, Fang-Okak is encircled in the North by Fang-Ntumu and Fang-Mvaï, in the South by Fang-Atsi, in the East by Fang-Nzaman and in the West by Fang-Mekè. Because of this position, Fang-Okak is the only dialect of Fang in Gabon that is in permanent contact with the rest of the

¹⁷ It is the sole province of this country where there is only one language. In the rest of the provinces, many languages are cohabiting together.

¹⁸ I did not find or even heard about any work made in this dialect.

other dialects. This means that Fang-Okak is the most influenced dialect of Fang. The differences between Fang-Okak and the other dialects (especially with Fang-Atsi, Fang-Mekè and Fang-Nzaman) are difficult to identify.

Map 2: Fang dialects distribution in Gabon



2.3. 7. Fang-Nzaman dialect

Fang-Nzaman is located in the regions of Makokou, Koumaméyong, Booué and Ovan. I inventoried two works on this Fang dialect, i.e. linguistic research done by Andeme Allogho (1980) and Cinnamon (1999). Both of them compiled phonological outlines, with some cultural and historical aspects of this dialect. The most recent work in Fang-Nzaman has been done by Mekina (1997). This work is a morpho-phonological study of terms that Fang-Nzaman borrowed from French.

2.3.8. General works in Fang

Apart from works compiled in different dialects, some general research has also been done in Fang. Bancel (1986) in *Etude comparée des langues du groupe fang pour la création d'une base de données Bantu*, did a comparative study between all the Fang dialects. The goal of this work was to include Fang data in a database of Bantu languages. One of the earliest works about Fang was produced by Bennet (1899), who established some *Ethnographical notes on the Fang*. Balados Carter (n.d.) produced in Spanish *Elementos de la gramática pamue*, which presents some grammatical elements of Fang. Chamberlin (n.d.), made a new interpretation concerning *The migration of fang into Central Gabon during the 19th century*. Echegaray (n.d.) tried to shape the harmonization of Fang spelling by proposing a typical orthographical system for the writing of this language in *Hacia la unificación ortografía de la lengua pamue*. With his publication *Introduction to Fang, a Bantoid*

language, Harding (1968) compiled a kind of introductory history for Fang. Some years later, Kelly (1974) did a research on *close vowels* in Fang. Ndongo Esono (1956), Osorio (n.d.) and Tardy (n.d) have also done some work on Fang: *Gramática pamue, Vocabulary of the fan language in Western Africa* and *Contribution à l'étude du folklore bantou: fables, devinettes et proverbes fang*, respectively. These works focus on grammar, vocabulary and proverbs, riddles and fables in Fang. In *Essai sur la phonologie panchronique des parlers fang du Gabon et ses implications historiques*, Medjo Mvé (1997) did a phonological study of Fang, by comparing all dialects and showing the implication of the history in the phonological differences of Fang. Raponda Walker foundation published the most recent works on Fang. The first one, *Rapidolangue* (1998) is a learning book presenting 5 Gabonese languages, e.g. Fang, Yipunu, Yinzédi, Lembaama, Omyénè, plus English and French. The second one, *Les langues du Gabon* (2000), is a comparative study between some Gabonese languages, including Fang.

2.4. Choice of the Standard dialect

2.4.1. General remarks

Standardization in any sphere of activity involves imposing uniformity upon a class of objects and suppressing variation¹⁹. That suppression of variation in language will ensure communication over longer distances of space and time with a minimum of misunderstanding.

¹⁹ The suppression of variation here does not mean that the other dialects should be suppressed or will not exist any longer when the standard dialect has been chosen. Suppression of variation means only one form or variety of the language will be used in all official activities (schools, administrations, etc.).

Standard languages get established as groups of speakers perceive the needs for a set of shared linguistic norms. Language variation derives importantly from the different uses to which a language is put (writing as opposed to speech, formal as opposed to informal contexts, planned, message-oriented discourse as opposed to unplanned, listener-oriented discourse, etc.), and receives such labels as 'model', 'style' and 'field' (Anthony Lodge, 1993:13). "Standardization involves the progressive elimination of alternative norms through the selection of one norm which is superimposed on the rest" (Anthony Lodge, 1993:24 & 95).

One of the easiest ways to uniform a language is to choose, among all its varieties, one of them as the standard language. That choice is not done in any method, it must follow some scientific procedures already experimented by other researchers. The purpose of the following section is to identify these criteria and see which one of the Fang dialects complies with most criteria.

2.4.2. Criteria

One of the aims of this research project is to devise a model for a standard translation dictionary, which could influence the standardization of Fang. The choice of a standard dialect is the first step in the standardization process of a language with several dialects. The model I am going to suggest also has to take into account the matter of dialects in Fang. One dialect from this language should be proposed as standard dialect. This should be done according to some well-defined criteria. Garvin & Mathiot (Gallardo,

1980) define a standard language (dialect) as *a codified form of a language, accepted by and serving as model to a large speech community*.

Zgusta (1980:170) reminds us that in the history of different languages, one can observe the frequent phenomenon that for various reasons, mainly of a cultural, political, and economical nature, one of the dialects of a language gains preponderance over the other dialects, or a good part of them. The preponderant dialect develops into the standard national language. This standard national language usually has some variants, e.g.:

- (1) The literary language and the cultivated spoken language.
- (2) Colloquial language is the form of the standard national language which is spoken in private and in semi-official situations (e.g. business) by people of average and higher education.
- (3) Folk speech is the form of the standard national language spoken by people of minimum or no education who do not, however speak a (rural) dialect.

Fromkin & Rodman (as cited by Anthonissen & Gough, 1998) affirm that the dominant or prestige dialect is often called the standard dialect. A standard dialect may have a social function: to bind people together or to provide a common written form for multidialectal speakers. Holmes (as cited by Anthonissen & Gough, 1998) adds that the standard variety is generally one which is written and which has

undergone some degree of regulation or codification (e.g. in grammar and in dictionaries).

For Gallardo (1980) the way in which the standardization process takes place will depend on the circumstances surrounding the development of each speech community. One of the most important elements that determine the standardizing trend is the nature of the language models that operate in each situation. For instance, religious, literary or technological language models are associated with different attitudes among the speakers, and will generate different ways to cultivate the language. Emejulu & Nzang-Bie (1999) argue that there is no specific scientific model for choosing a standard dialect. According to them, Sadembouo (1980) listed 18 criteria for selecting a standard dialect that he classed under three headings: **fundamental**, **secondary** and **marginal** levels:

- **Fundamental criteria**

- High degree of declared understanding of the dialect.
- High degree of predicted understanding of the dialect.
- Numerical importance of the dialect speakers.
- Advantageous geographical position of the dialect.
- The location of the dialect at the center of activity.
- Dialect prestige.
- Pureness of the dialect.
- Mobility of the dialect.

- **Secondary criteria**

- The attitude of the government towards the dialect.
- Religious influence of the dialect.

-
- Socio-economic importance of the dialect.
 - Written documents already existing in the dialect.
 - Historical expansion of the language.
 - Expressed feeling on the ease of understanding and speaking of the dialect.
- **Marginal criteria**
 - Availability of dialect speakers ready to cooperate in the language development work.
 - Good working conditions for the researchers.
 - Friendship relations between the researchers and a speaker of the dialect.
 - Social status of the dialect speakers.

To these criteria proposed by Sadembouo, Emejulu & Nzang-Bie (1999) add the following aspects:

- National languages of wider communication.
- Decentralization of decision of choice of language to regions.
- Dominate regional languages.
- Choice left to the communities.
- Native tongue of the capital city.
- Presence of leaders engaged in the standardization process.
- Presence of a committee concerned with standardization.
- Endangered languages (dialects). This criterion will serve in preserving these languages. John de Vries (1991: 51) cited the case of Swedish in Finland, which has been protected by a language law dated to the early years of the 20th century.

This Swedish example can be taken as a model for protecting endangered languages (dialects).

- The mutual comprehension rate between dialects.
- Accepting the standardization of a second dialect if the dialects of the language are structurally wide apart with minimal mutual comprehension.

For Emejulu & Nzang-Bie (1999), the most important and more workable criterion is the **acceptability of the language community**. They are referring to Sadembouo, Tadadjeu and Wiesmann (1988), who declared, "*People decide, not principles*". Emejulu and Nzang-Bie (1999) confirm that most languages have many varieties or dialects, which are spoken in different geographic areas. A reference dialect serves to unify the language. The aim is not to make all members of the community speak in the same way; but through the written variety language unity will develop around the standard dialect.

The functions of the dialect in the speech community have been described by Anthony Lodge (1993:118-152) as one of the major criteria as far as the choice of the standard dialect is concerned. Indeed, he estimates that a standard dialect could be chosen according to its sociolinguistic functions, e.g. if the dialect is used in a wide range of official and public functions, not only as the language of the government and the law but also as a language of learning (Anthony Lodge, 1993:118). Another important characteristic in the choice of the standard dialect can be the prestige. Prestige here means the language of power in the community usually perceived to

be superior, more elegant, more beautiful and more logical than the other varieties (Anthony Lodge, 1993:130).

Anthony Lodge has retained six main criteria that can help in the choice of the standard dialect as presented in the following table:

Table 4: Criteria, role and applicable fields

| Criteria | Role or applicable fields |
|----------------------|--|
| 1. Functions | Religion, conversation, education/learning, administration, instruction to servants, real literature, folk literature. |
| 2. Prestige | When the dialect is felt to be more beautiful, more logical, better able to express important thoughts, has a strong association with religion |
| 3. Literary heritage | If literary documents already exist in the dialect. |
| 4. Acquisition | If the dialect has been taught (in schools for instance). |
| 5. Standardisation | If there is a unique form of the dialect and that has been codified (e.g. grammar) |
| 6. Stability | If the dialect has been used for several centuries. |

After an overview of criteria stressed by researchers, the following table shows how each Fang dialect meets with the inventoried main criteria: the black dot indicates that the dialect fits the corresponding criterion.

Table 5: Correspondences criteria/dialects

| Criteria \ Dialects | Fang-Atsi | Fang-Mekè | Fang-Mvāi | Fang-Ntumu | Fang-Nzaman | Fang-Okak |
|--|-----------|-----------|-----------|------------|-------------|-----------|
| 1. Codified form | | | | | | |
| 2. Accepted form | | | | | | |
| 3. Serving as model | | • | | | | |
| 4. Binding people together | | | | • | | |
| 5. Providing a common written form | | | | | | |
| 6. Having a written form | • | • | • | • | • | |
| 7. Model operating in religious language | | • | | | | |
| 8. Model operating in literary language | • | • | • | • | • | • |
| 9. Model operating in technological language | | | | | | |
| 10. High degree of declared understanding | • | • | • | • | • | • |

| | | | | | | |
|---|---|---|---|---|---|---|
| 11. Numerical importance of speakers | | | | • | | |
| 12. Advantageous geographical position | | • | | • | | |
| 13. Location of the dialect at the center of activities | | | | • | | |
| 14. Purenness | | | • | | | |
| 15. Socio-economic importance | | | | • | | |
| 16. Written document already existing | • | • | • | • | • | |
| 17. Historical expansion | | • | | • | | |
| 18. Availability of dialect speakers to cooperate in the language development | • | • | • | • | • | • |
| 19. Good working conditions for the researchers | | • | | • | | |
| 20. Dominate regional language | | | | • | | |
| 21. Native tongue of the capital city | | | | • | | |

Statistical results: The following results give the number of criteria from the 21 used which each dialect fits.

Fang-Atsi: 5

Fang-Mekè: 10

Fang-Mvaï: 5

Fang-Ntumu: 14

Fang-Nzaman: 5

Fang-Okak: 3

2.4.3. Comments

For a better understanding of the results, some comments will be necessary. The reason why each dialect fits with any criterion or not can be understood by looking carefully at the sociolinguistic and historical role of each Fang dialect in the Gabonese context:

1. At the current stage of the development of Fang, no dialect has been codified yet. Many linguistic work have been undertaken (e.g. phonological, morphological, syntactical descriptions, etc.), but those works still remain purely scientific, only used by linguistic experts and unknown to laypersons (e.g. common speakers).
2. Despite the fact that some Fang dialects have been used in common contexts (e.g. the religious field) by speakers from different dialects, none of the Fang dialects have imposed itself on the other dialects. People are still strongly rejecting the other dialects in order to see their dialect imposed on the others.
3. Until now, Fang-Mekè is the sole Fang dialect that has really served as model because of its particular role in the religious field. Indeed, that dialect has been used for the translation of religious texts (e.g. the Bible, catechism, etc.).
4. Any Fang speaker from the Woleu-Ntem province whatever dialect they speak will easily recognise that Fang-Ntumu is the main dialect used for all socio-economic and cultural activities in that province. People from different parts of the province are going to Oyem and Bitam (which are Fang-Ntumu areas) for different reasons (e.g. working, learning, shopping, etc). Most of these activities are taking place in Fang-Ntumu. For that reason this dialect is binding Fang people together in the Woleu-Ntem province.

5. Despite the fact that most Fang dialects have been described, most of the written forms in these dialects are compiled in phonetic transcription. However, there are also existing lexicographical works compiled in different Fang dialects. In these works, different writing systems have been used, even within one dialect. For that reason there exists no common written form in any Fang dialect.

6. Many works have been compiled in different Fang dialects. Therefore, each dialect has, at least, a written form.

7. The point here is not to prove that one dialect can operate in religious language more than the others (any dialect can do so), but to identify the dialect that has played the most important role within the religion. In that regard, Fang-Mekè cannot be ignored because most of the existing religious texts in Fang are in that dialect. That situation can be explained by the fact that Fang-Mekè is spoken in Libreville, capital city of Gabon and siege of the clergy. In addition, André François Ndong, the first Gabonese Archbishop, who compiled many religious works in Fang, spoke Fang-Mekè.

8. Fang, including its dialects, operate in literary language. Unfortunately the literary role of Fang remains on the oral level. Indeed, most of the literary works based on Fang culture are compiled in French because of the lack of a common written form of Fang.

9. One of the reasons that explain the lack of technological terms in Fang is the fact that this language is not yet used in the academic field. Indeed, these technological terms will be generated when the user feels the need to name them, particularly for academic or scientific purposes. The process has already started in the Gabonese media (public and private), particularly radios (e.g. RTG²⁰, Radio Soleil²¹, etc.) by the organization of programmes where participants are trying to translate, different words and concepts used in scientific and/or academic fields (e.g. mathematics, physics, biology, etc.) in different local languages.

10. The biggest majority of Fang speakers whatever their dialect easily recognise the high degree of understanding between all the Fang dialects.

11. Fang-Ntumu speakers are numerically more than those from the other dialects.

12. Fang-Mekè (Libreville) and Fang-Ntumu (Oyem) by the fact that they are located in capital cities are geographically well positioned. That geographical position gives them the advantage to be easily reached by people from other parts of the country.

²⁰ RTG (Radio Télévision Gabonese) is the first public broadcasting channel in Gabon.

²¹ Radio Soleil is a private Radio station.

13. In Woleu-Ntem, the sole monolingual province of Gabon, Fang-Ntumu is the primary dialect. For that reason it is at the center of most activities between Fang people.

14. Of all the Fang dialects, Fang-Mvaï is sometimes designated as Fang-Fang²², in order to qualify it as the purest or the one closer to the original Fang. However, that hypothesis has to be taken with caution, because that theory has not been proven yet by any research.

15. By the fact that it is at the center of activities in Woleu-Ntem province, Fang-Ntumu is playing an influential role in socio-economic activities.

16. Apart from Fang-Okak, the Fang dialect in which few works have been compiled, all the Fang dialects do have existing written documents. Most of these documents are linguistic and lexicographical works compiled by missionaries.

17. The history of the Fang people shows that Fang-Mekè speakers have been very active in their migration through Gabon. They have tried to reach and conquer new areas in the country and to influence the use of their language everywhere they have settled. As far as Fang-Ntumu speakers are concerned their expansion has been more visible in the socio-economic plan. Indeed, it is well recognised that speakers from that dialect are hard workers and successful people. That is illustrated by the high number of educated people from that

²² Fang from the original Fang

dialect and the key positions many of them occupy in Gabonese society.

18. It is obvious that all speakers from all dialects would like to see the development of their dialect. Therefore they will do everything to help with the development of the language. Their cooperation in that regard is a certainty.

19. Most lexicographical research necessitates the presence of lexicographers in the field. For that reason the possibility or not to easily reach an area can be perceived as a criterion for the selection of the standard dialect. Contrary to the other dialects, Fang-Mekè and Fang-Ntumu located in areas (Libreville Oyem and Bitam) are easy to reach and have better working conditions for the researchers (e.g. roads, telephone, hotels, libraries, etc.).

20. As already said, Fang-Ntumu is the most frequently used dialect in the Woleu-Ntem province. For that reason it can be considered as the dominant language of that region.

21. Oyem, one of the original areas of Fang-Ntumu, is the capital city of the Woleu-Ntem province. Therefore, it is the native tongue of the capital city. The same situation could be applied for Fang-Mekè, which is spoken in Libreville, capital city of Gabon. But the influence of that dialect is not really visible in that area, because of its cohabitation with French, many other Gabonese languages and Fang dialects.

2.4.4. The proposed standard dialect

After looking at all criteria guiding the choice of the standard dialect, I propose Fang-Ntumu as standard model because of the following reasons:

- Fang-Ntumu is the dominant dialect in Fang because of the number of its speakers.
- Fang-Ntumu plays an important economic, cultural and social role in the province of Woleu-Ntem. The main activities in this province are taking place in Oyem and Bitam, which are original regions of Fang-Ntumu. A lot of people from the others parts of this province and from neighbouring countries (Cameroon & Equatorial Guinea) are going there to work, study and do business. Fang-Ntumu is the language of communication between the Fang population in this part of Gabon.
- From all the varieties of Fang, works published in Fang-Ntumu are more accessible to people today. Most of the works in the other dialects are really not available in Gabon. For instance, dictionaries compiled in Fang-Atsi and Fang-Mekè cannot be found in any library in Gabon.
- Fang-Ntumu is the Fang-dialect mostly used in the audiovisual media (radio & television) in Gabon. Fang-Ntumu

also is the first dialect of Fang in which an entire movie has been made.

All these criteria have motivated the proposal of Fang-Ntumu as dialect of reference. Even if the community can choose itself²³ the standard model, as proposed by Emejulu and Nzang-Bie, the lexicographer, by using one dialect, could guide this choice and accelerate the language standardization process.

2.5. Concluding remarks

As this chapter shows, a lot of work has been done in Fang. Three Fang dialects, Fang-Atsi, Fang-Mekè and Fang-Ntumu are those in which most works have been compiled. But the largest part of these works is either general or directed at linguistic purposes. Any user cannot access them.

The main lexicographical works in Fang are from Lejeune, Martrou, Largeau and Galley. Most of those dictionaries and encyclopaedia were compiled by non-lexicographers, without metalexical criteria. Those works are also not available to the majority of the people and rather had political and anthropological objectives. No specific spelling was used in those works and they are not easy to read, even by Fang mother tongue speakers. Those works show the necessity to shape a new way to compile dictionaries in Gabonese languages in general and Fang in particular. This new way will not only facilitate the making of bilingual dictionaries in Gabonese

²³ The option where the community itself choose the standard model can take so much time, because everyone will try to impose his dialect.

languages, but will also help to standardize them. The compilation of a standard dictionary in a non-standardized language is always a big matter that the lexicographer has to solve. The usual way followed by many lexicographers is the choice of one dialect as the standard one. But this choice is not easy and still poses some controversial issues, even between researchers (lexicographers and linguists):

- Some researchers estimate that the choice of the standard dialect should be left to the people, because they are in a better situation (as speakers) to know which variety of the language could play the role of reference in the language use.
- The others think that letting the choice of the standard dialect to the people will delay the standardization process, because everyone would propose his dialect (even his idiolect) as the standard one. The people do not care about the scientific criteria, they just want to be sure that their language (dialect) will not (at least) disappear or will dominate (why not) the others.

Personally, I agree with those who think that the lexicographer should propose and use one dialect as the standard one according to some scientific criteria. Those criteria are coming from some practical experience in the standardization of languages. It is sure that whatever dialect the lexicographer will propose, it will not immediately be accepted by the speakers from the other dialects. But by using scientific criteria the lexicographer should logically make a

right choice that the people will accept after a certain time. Indeed, when a language variety is used in a dictionary, a large number of users will utilize it (the dictionary) without paying attention to the dialect it uses. That will accelerate the standardization process. It is important to note that the choice of the standard dialect could be made in several ways (e.g. governmental decision as done in many countries in Europe). I believe that the lexicographer's choice is the right way, as I will demonstrate in the chapter concerning the role of the dictionary in the standardization process. Among all the dialects of Fang, any of the six varieties could be chosen as the standard one. My choice went to Fang-Ntumu because it seems to be, in my opinion, the one which is closest to the criteria I inventoried. Despite those criteria, I still consider that this choice is subjective as any other choice of the standard dialect even by the people themselves.

CHAPTER 3:

Dictionaries and corpora

If African linguistics is to take its rightful place in the new millennium, the active compilation, querying and application of corpora should become an absolute priority (Prinsloo, 2002).

A frank epistemological rupture is necessary to make the study of Gabonese languages a useful practice, and has the development one expects from them. To reach this goal the linguists and lexicographers of Gabon must preferentially, from now on, direct their resources and efforts toward the compilation and development of electronic corpora (Emejulu, 2001; translated from French).

3.1. Introduction

A corpus can briefly be described as a very large collection or a body of words, usually stored in computer format. The word *corpus* is from the Latin word for "body". The corpus plays an important role in the dictionary compilation process. A corpus is designated according to some criteria and drawn from specific fields. In the language sciences *a corpus is a body of written text or transcribed speech, which can serve as basis for linguistic analysis and description*. Linguistic corpora are not ends in themselves, but they are rather one source of evidence for improving descriptions of the structure and use of languages, and for various applications, including the processing of natural language by machine and understanding how to learn or teach a language (Kennedy, 1998:1). In language study, *a corpus is any body of text collected with the aim of analysing its features*. Nowadays when one speaks of a corpus in lexicography, it refers to an electronic corpus, often containing a vast number of words from

many different sources (Landau, 2001: 273). A corpus constitutes an empirical basis not only for identifying the elements and structural patterns which make up the systems we use in a language, but also for mapping out our use of these systems. A corpus can be analysed and compared with other corpora or parts of corpora to study variation. Most importantly, it can be analysed distributionally to show how often particular phonological, lexical, grammatical, discursal or pragmatic features occur, and also where they occur (Kennedy, 1998:4). Lexicography is usually thought of as covering the collection of lexical items and a description of the way they are used. Typical lexicographical products include not only printed dictionaries of various types, but also concordances, indexes, terminologies, and so on (Zampolli, 1994:4). The main goal of this chapter is to review some practical stages in the process of corpus building, from the selection of sources through to mark-up and, assigning annotation to assembled texts (Atkins, Clear & Ostler, 1992). Particular attention will be paid to computational methods and tools designed to assist the various lexicographical tasks, including the preparation of lexicographical evidence from many sources, the recording in the corpus form of the relevant linguistic information. Mention will also be made of the population group from which the data (oral & written) will come, with the objective to compile a standard French/Fang bilingual dictionary. An important discussion related to the role of data processing (computers, software and networks) in the compilation of a corpus will take place in the following sections. Prinsloo (2000) considers that compiling and

querying electronic corpora has become a *sine qua non* as an empirical basis for contemporary linguistic research. He argues that good dictionaries increasingly base the compilation of both their macro- and microstructure on electronic corpora.

By referring to Davidson, Heung-Yeung & Deerwester (1994:12) and Calzolari (1996:4), Prinsloo (2000) specifies that the crucial benefit of corpora in the broad field of linguistics has been summarized as follows in the literature:

- Corpora are the research tools for the essential inductive observation of “real language”
- Presently, corpora are recognized, by more and more research and development groups, as the most precious aid in designing systems that respond to user needs, in terms of types of texts and real language to be treated.

Many approaches and experiences have been described for the compilation of a corpus in the lexicographical field. Approaches and experiences about corpus building in African languages, particularly from Prinsloo (2000), Engwall (1994), Atkins, Clear & Ostler (1992) will be the main routes used in my proposed model. I will also use some general theories in corpus building, e.g. Kennedy (1999), Gouws (2001), Béjoint (2000) and Landau (2001). Finally, for each step I will propose my own point of view as far as the compilation of a corpus for a French/Fang standard bilingual dictionary is concerned.

3.2. Corpora typology

There are different types of corpora. Each type is compiled according to the objectives of the compiler. From these corpora, I have inventoried 5. The specificity and the usefulness of each one in the lexicographical field are highlighted:

- **Written Corpora:** Written corpora are essentially written texts that enable the lexicographer to find many of the possible meanings and uses of individual words, and also frequency information about them. For example, in a dictionary, the word "access" occurs as a noun and a verb. By using a written corpus, the lexicographer could also note that the noun form is much more commonly found than the verb form. This could be a useful piece of information for language teachers, or for translators. The lexicographer can also compare different written genres - how does academic writing differ from journalism? How does the written language of 50 years ago compare to the language used today?
- **Spoken Corpora:** Corpora of spoken language data help the lexicographer to look for discourse patterns in language, e.g. different ways that people manage arguments and disagreements. If information about each speaker is included, then more complex questions could be asked, e.g. do men talk differently to women? Spoken corpora could also be used to tell the lexicographer about accents and dialects in different parts of the investigated area.

-
- **Parallel Corpora:** These consist of the same texts, translated into different languages. For example, a corpus could contain the same texts in Fang, Afrikaans, English, and French. This type of corpus is useful for developing ways of automated translation from one language to another, or in the creation of bilingual dictionaries, or in language teaching.
 - **Learner Corpora:** Learner corpora mainly contain extracts of language that have been produced by people who are learning a new language. This type of corpus can give the lexicographer much information about how the language develops. The lexicographer can also use a learner corpus to collate mistakes that learners are most likely to make, and this can help the improvement of the teaching of that language. For example, one researcher looking at learner corpora has found that one of the most common mistakes for learners of English was that the word "people" often tends to be written as "peoples". As a result more specific instructions on the use of "people" were incorporated into an English-language textbook.
 - **Specific Corpora:** A corpus can also focus on a more specific form or genre, particularly forms of language that are not studied in much detail very often, in order to make sense of how language is used in that context. For example, it is possible to compile and analyse a corpus of stories for children, a corpus of internet chat-room discussions or a corpus of paranormal stories. Specific corpora can also serve for the compilation of LSP dictionaries (e.g. medicine, mathematics, etc.).

The corpus that needs to be compiled for the proposed dictionary model will be a kind of hybrid, because it will include some features from the presented corpora:

- It will be a written corpus because it will include written data from sources such as dictionaries, newspapers, glossaries, religious documents, etc. Written sources will enable the compiler of a French/Fang standard translation dictionary to find many of the possible meanings and uses of individual French and Fang words, and frequency information about them. This information could be useful for the lexicographer when looking for translation equivalents in Fang. The written sources will also allow the lexicographer to compare the way French and Fang are written in different situations. As Fang does not have many written documents, the use of written sources will focus more in French.
- It will also be a spoken corpus because it will include discourses and conversations in French and Fang. Spoken sources will allow the compiler of the French/Fang standard translation dictionary to identify different ways that Fang people manage arguments and disagreements. They will also allow the lexicographer to have a better idea of the differences between dialects. The use of spoken material will be more important in Fang as it is a language that has few written documents.

-
- As the proposed dictionary will be a bilingual one, the intended corpus for a French/Fang standard translation dictionary should also be a parallel corpus. Indeed, all texts compiled in French will be translated into Fang and vice versa. This corpus should enable Fang lexicographers not only to compile French/Fang bilingual dictionaries, but also to develop ways of automated translation from French to Fang, and vice versa.
 - The intended corpus will also be a learner one since it will contain extracts of French and Fang produced by the target group. Thanks to these extracts, Fang lexicographers will be able to collate mistakes that young Fang learners are likely to make in both languages. Therefore, lexicographers could give more specific instructions on misused words in dictionaries.
 - Finally, the intended corpus will also be a specific one since it will include data from some specific sources, particularly those from academic fields. This corpus will not only serve for the inclusion of LSP terms in the French/Fang standard translation dictionary, but also for the compilation of LSP dictionaries.

In the following sections I discuss the compilation process of a corpus for a French/Fang standard translation dictionary. As the proposed dictionary model is a bilingual one, the described compilation process has to be done in both French and Fang.

3.3. The material

3.3.1. Introduction

The dictionary compilation presupposes a well-devised plan. In the planning of a dictionary project, the dictionary conceptualisation plan has the most direct influence on the compilation process (Gouws, 2001:67).

According to Wiegand (Gouws, 2001: 67) the dictionary conceptualisation plan can be divided into five subdivisions:

- General preparation phase.
- The material acquisition phase.
- The material preparation phase.
- The material processing phase.
- The publishing preparation phase.

Even if the general preparation phase of the dictionary conceptualisation plan lays the foundation for the structure, contents and presentation of the final product (Gouws, 2001:97), the present section will primarily be focusing on two of the five phases cited by Wiegand and summarized by Gouws (2001:69-70), particularly those ones related to the collection and preparation phases of the data: **the material acquisition** and **the material preparation**.

3.3.2. Material acquisition

3.3.2.1. General approaches

A logical phase to follow the identification and formulation of a dictionary basis in the general preparation phase of the dictionary conceptualisation plan, is the material acquisition phase. This phase

precedes the compilation process. The material acquisition phase focuses on the gathering of speech material from the sources earmarked for the dictionary basis (Gouws, 2001:69). The first stage in compiling a corpus is to decide why it is needed and what its purpose will be. The nature of the corpus will depend on the size and scope of the dictionaries for which it is to be designed, its intended audiences, and the resources (in both people and money) that can be marshalled for the effort of compilation (Landau, 2001: 323). The first responsibility for the lexicographical team is to determine the socio-cultural level of the dictionary it wants to compile, and establish according to this level, the vocabulary (terminology) and information about this vocabulary (Dubois & Dubois, 1971:18). This cultural level, which involves the linguistic needs of the target-group, will determine the type of material that should be collected. The material collection should be done where this target group exerts its activities.

The two fundamental constraints for the construction of any linguistic corpus are the *objective of the study* and the *resources available* (Gunnel Engwall, 1994:52). The dictionary basis of a general or bilingual dictionary can be compiled from three types of sources; the primary sources, the secondary sources and the tertiary sources. A result of the material acquisition phase is the compilation of the lexicographic corpus – the collection of items gathered from the primary, secondary and tertiary sources of the dictionary basis (Gouws, 2001:68-69).

3.3.2.2. Criteria for the material selection

Svensén (1993:40) says that it is necessary to decide before the material acquisition, the sources from which this material will be retrieved. The selection of these sources has to be done according to some criteria. In my view, the following criteria are the most important as far as the selection of sources of a French/Fang bilingual dictionary is concerned:

- **Authenticity.** The lexicographer must ensure that the selected sources are giving the current use of the language and that those sources are credible enough. In the case of oral sources, the lexicographer must be sure that his sources reflect the language as used by the target users.
- **Coverage.** The lexicographer has to make sure that the sources cover the largest possible range within the area of the language to be described.
- **Suitability.** Sources should only be used if the lexicographer is sure that they will give him/her useful data. That source must be available to allow the lexicographer to retrieve the data he/she needs.
- **Stylistic aspect.** Sources must reflect all styles in the use of French (high, medium and low levels of the language use) and these must then come from different areas where the target group is using French.
- **Time aspect.** Sources must be selected according to their synchronic features as well as the diachronic ones. In fact, the

compilation of the corpus has to take into account old and new words.

For Bo Svensén, the collection of data must also observe some primary rules or principles:

- The main objective is the collection of data, not evaluation or editing. Avoid unnecessary comments on the words and do not suggest definitions or translation unless these arise “spontaneously”. Conscious searching for definitions or translations at this stage diverts the attention from the real task.
- Excerpts must not be too long or subjected to the demands of intelligibility and suitability. Omissions should be marked.
- It is extremely important that the source should be stated. An excerpt whose origin cannot be traced is of very little use.
- Do not take the same word from the same source and with the same meaning more than a certain number of times (the number must be decided beforehand).

By following all those criteria and principles, the lexicographer can be sure that the selected sources will give him the right data and give him the opportunity to reach his objectives as far as the compilation of the corpus for a French/Fang bilingual dictionary is concerned.

3.3.2.3. Primary sources

The primary sources of the dictionary’s basis will be all the written material reflecting typical communication situations. Although the

primary sources will usually be texts, the dictionary basis of a dictionary compiled for a language with a strong oral tradition can also use recordings of the orature as primary sources (Gouws, 2001:68). If printed and electronic sources are limited, there is no way around it: one has to record a large variety of spoken speech from as many different genre/topic areas as possible (Prinsloo, 2000).

By using the structuralist and generativist approach in which the linguistic system should be studied as a whole, Gunnell Engwall (1994:50) argues that the linguistic system, or "langue" (language) as Saussure termed it, is virtual because it cannot be studied directly. The only way to study this system is by its actualisation (expression): the word ("parole" according to Saussure terminology). Gunnell Engwall adds that "Parole" may be thought of as all actualised utterances, written or spoken, up to the start of the study, and constitutes the total population. Since its size precludes full-scale studies, as linguists (lexicographers in our case) we are obliged to select parts of the population as objects of our research. This selection may be made by means of chance or choice. The chance means that researchers simply collect examples when reading books or newspapers or listening to conversation.

Researchers who select their material in an organized way must adhere to a careful definition of the specific subpopulation of the natural language that is the object of their study. If each domain cannot be defined by consensus, clear borderlines should be established, in order to allow the others to fully identify text

categories involved and criteria for sampling. Selection based on deliberate choice has therefore several advantages for corpus construction.

The compilation of the oral sources has to be, as far as the compilation of a French/Fang bilingual dictionary is concerned, the most important stage of the data collection, because the objective is not to collect any French speech, but that of the young Fang speaker in Gabon. In this regard the oral speech from the target group must constitute the biggest part of the corpus. The material should be collected by recording speech where young Fang speakers currently use French, e.g. schools, areas of entertainments (sports and games), ceremonies, night clubs, meetings, parties, churches, media (TV, radio, newspapers and magazines), as well as any available audio or video tape. Kennedy (1999:4) strengthens this approach by saying that a corpus constitutes an empirical basis not only for identifying the elements and the structural patterns which make up the systems we use in a language, but also for mapping out our use of these systems.

A corpus can be analysed and compared with other corpora or parts of corpora to study variation. Most importantly, it can be analysed distributionally to show how often particular phonological, lexical, grammatical, discursal or pragmatic features occur, and also where they occur. This kind of information can be better obtained from a corpus compiled with oral sources than one compiled with written sources (particularly non discursive written sources).

The collection of this data must cover all Fang areas in Gabon. The proposed dictionary is a standard one, it does not only concern speakers from the standard dialect but all young Fang speakers throughout the country. The lexicographer has to be sure that the collection of data takes into account every niche where young Fang people speak French. Even if the majority of the Gabonese population live in the major towns, the collection must not only be carried out there, but also in smaller areas that has intensive human activity and where the target group is linguistically active in French, particularly where there are schools, academic institutions, commercial activities, hospitals, guest houses, etc. The lexicographer's investigation must be done in all areas where there is any variety of Fang. In Gabon, the way Fang people speak French is much influenced by their Fang dialect and by the region in which they live or come from. All the varieties of French used by Fang speakers have to be taken into account. In some cases the dialect spoken in the same area can be different from one part to another. In Minvoul for instance, the Fang-Mvaï dialect can be divided into three varieties: Fang-Mvaï spoken in the north, close to Cameroon, Fang-Mvaï spoken in the centre and Fang-Mvaï spoken in the south, close to Oyem. The same situation can be observed in the Fang-Ntumu dialect. The way this dialect is spoken in Oyem is different from the one spoken in Bitam. All those particularities must be taken into account by the lexicographer during the recording phase.

The collection of oral sources in the Estuaire region should be done carefully. More than half of the Gabonese population live in this

region (particularly in Libreville). This means that it is the most linguistically mixed region of the country where all dialects (and languages) are represented. The investigation in this area must therefore be carried out with a particular diligence. Indeed, the original region of the investigated group (or person) has to be distinctly noted. Such precautions are necessary for the subsequent data analysis and lexicographical treatment. The lexicographer will need to know in which region and/or group each lemma and/or item equivalent is mostly, or exclusively, used. It is important to remember that the linguistic needs of the target group should be taken into account in the collection of data. For instance, young Fang speakers will probably need to learn the study subjects in French at school in Fang. Terms related to mathematics, biology, chemistry, geography, economics, computing, etc. should be included in the proposed dictionary model. One of the ways to reach this goal could be the organisation, by the lexicographer himself, of discussions or debates about themes or topics related to the concerned subjects. These informal discussions and debates can be recorded by the lexicographer and then added to the corpus. That will allow the lexicographer to compile a corpus, which takes into consideration all the linguistic needs of the target group.

The method of informal discussions also called “production libre” (free production) by Nzang-Bie (2002) should be done with speakers that are competent enough in the language. According to the experience of Nzang-Bie in the collection of oral data in Fang-Mekè, “free production” means that each language user should be

free to develop his thought in the proposed theme and anyone can intervene or say something at anytime when he wants to react to what another said. Nzang-Bie (2002: 218), who used that technique to collect oral data in Fang-Mekè said that the method has certainly not been used by many researchers up to now, but if used, it should give evident results for the compilation of future lexicographical corpora.

According to Svensén (1993:54) primary sources can also be regarded as supplementary material to complement the material which constitutes the bulk of the dictionary and has been collected in different ways. For example, it may have been concluded that the existing material has shortcomings as regards:

- Words which appeared only recently;
- New or overlooked meanings of words already included;
- Slang expressions, colloquialisms, professional jargon;
- Technical terms;
- Variant spellings and inflexions;
- Constructions;
- Idioms.

For Svensén, the way the material is collected is often that a number of strategically chosen texts are examined, and from them "candidates" are selected for inclusion in the dictionary. When choosing texts it is important to make sure of a sufficient variety and extent, so that all variants of the language are represented.

3.3.2.4. Secondary sources

The secondary sources are all the available dictionaries in the specific language. In this regard lexicographers have to be very careful not to perpetuate lexicographic failures of the past in a new dictionary project (Gouws, 2001: 68-69). As far as the collection of material goes, the bilingual lexicographer is in an enviable situation if there is already a good, comprehensive, descriptive monolingual dictionary, preferably of the standard-descriptive or overall-descriptive type with, at least, a standard-descriptive nucleus of the non-native language of his pair of languages. It is even better if there are such dictionaries of both of these types (cf. Zgusta 1971:307). The absence of such a dictionary is always a serious handicap, because the lexicographer himself must then do much descriptive and other work that should have been covered by the monolingual dictionary. This refers not only to the collection of the material but has more general validity: in the absence of a monolingual dictionary, the lexicographer will not only have to decide for himself what items are to be considered as stabilized lexical units and what are not, but he/she will also have to deal with the multiple meanings of each lexical unit, etc. All this will be more difficult and take longer. As far as the collection of material goes a good monolingual dictionary can be used as the basis for the planned bilingual dictionary (Zgusta, 1971:307).

Even though the target group of the dictionary model I propose is young Fang speakers, the lack of a monolingual Fang dictionary should not be a real problem as far as the collection of material is

concerned, as Fang is the target language of the proposed dictionary model. The collection of data will primarily concern the source language, i.e. French. As already stated, the target group usually performs better in French than in Fang. The main goal of the proposed dictionary is to give the target group the opportunity to transfer its knowledge from French to Fang. As such, all monolingual French dictionaries are important sources from which data must be retrieved. Particularly all types of French monolingual dictionaries relevant to the young French speaking people (e.g. school and learner's dictionaries). Important information should also be retrieved from the existing Fang bilingual dictionaries referred to in chapter II. In this regard, the *Encyclopédie pahouine* of Largeau (1901) and the *Dictionnaire fang-français et français-fang* Galley (1964) constitute important sources from where information could be retrieved. According to Svensén (1993:58) secondary sources can serve several purposes. They can be used as a check on the main source. Therefore, one should make sure that secondary sources and main sources have been compiled independently of each other, and they should be consulted continuously during the course of the work.

3.3.2.5. Tertiary sources

The tertiary sources comprise all other linguistic material that can be used, e.g. linguistic monographs and papers and grammars (Gouws, 2001:69). Gunnell Engwall (1994:51-52) has stressed some procedures that should be followed in the selection of text for a corpus:

-
- The first step relates to the identification of the category of texts, such as literary or scholarly works, newspaper, or conversation.
 - The second step is the choice of genre (or genres and subgenres), for example imaginative prose, drama, scientific texts, or dialogue.
 - The third step is to determine the time, more exactly the period, in order to place the research temporally, either as a diachronic (or broad) or synchronic (or narrow) study.
 - The final step is to decide on, first, the individual texts, and then the actual sample of passages from these texts. If the entire text is not to be included, the sample passages must be selected, preferably using random sampling techniques. It should be noted that such sampling is never of individual words but rather of longer or shorter sections of text, which together will constitute the corpus.

The procedure outlined by Gunnel Engwall could be a good organized way to retrieve data from secondary sources. This data could be complemented by some written sources from the target group, e.g. letters, diaries, school tests (particularly essays or compositions), French books for the young that teenagers or children use in schools (e.g. story books). The purpose is to take into account the linguistic needs of all categories of young Fang speakers (from 15 to 35 years old) and to recover maximum data that represents the French used by the target group. All these sources should be complemented by any

available work compiled in Fang, particularly all linguistic and lexicographical works inventoried in chapter 2, e.g. Ondo Mebiame (1992), Mba Nkoghé (1982), Voltz (1989), Andeme Allogho (1980), Ndong Menini (1980). The linguistic work *De la phonologie à la morphologie du fang-ntumu parlé à Abomezok (langue Bantu A78)* compiled by Ondo Mebiame (1992) represents an extremely important source in this regard. This exceptional linguistic analysis, the biggest ever made in Fang, is a huge source which contains a large sample of data. More than 2000 Fang words are presented with their translation equivalents in French. According to Prinsloo (2000), the idea of using “document types” was introduced by Rundell, and has been defined as texts from a particular subject area, that have a cluster of relatively objectively identifiable features such as time, region, and medium. He specifies 10 such broad subject areas:

- Natural and pure science.
- Applied science.
- Social science.
- World affairs.
- Commerce and finance.
- Arts.
- Belief & thought.
- Leisure.
- Fiction.
- Poetry, drama, humor.

For the data collection in the model I propose, the same types of areas could be used, because they belong to the fields in which the

target group exerts or will exert its activities in French. If they want to perform the same activities in Fang, they will need to transfer their French linguistic competences to Fang.

3.3.2.6. Some remarks about text collection

One fundamental decision is to determine how much of the corpus should be spoken and how much should be written language (Landau, 2001:324). For the theoretical model I propose, I suggest that the spoken material should be dominant in the corpus, as there is less available written material from the target group. Furthermore, the purpose of the proposed dictionary is to give the target group the opportunity to transfer their knowledge from French to Fang. But because of the difficulties and expense of obtaining spontaneous speech, much spoken material consists of transcriptions of radio or television talk shows and political interviews (Landau, 2001:324). These types of collections¹ can also be used in the Gabonese context by recording TV and radio programmes presented by speakers from the target group.

Atkins, Clear & Ostler (1992:2) stressed some examples of units of speech which might be considered to be spoken texts:

- An informal face-to-face conversation.
- A telephone conversation.
- A lecture.
- A meeting.
- An interview.

¹ Landau calls these types of speech *semiscripted*.

- A debate.

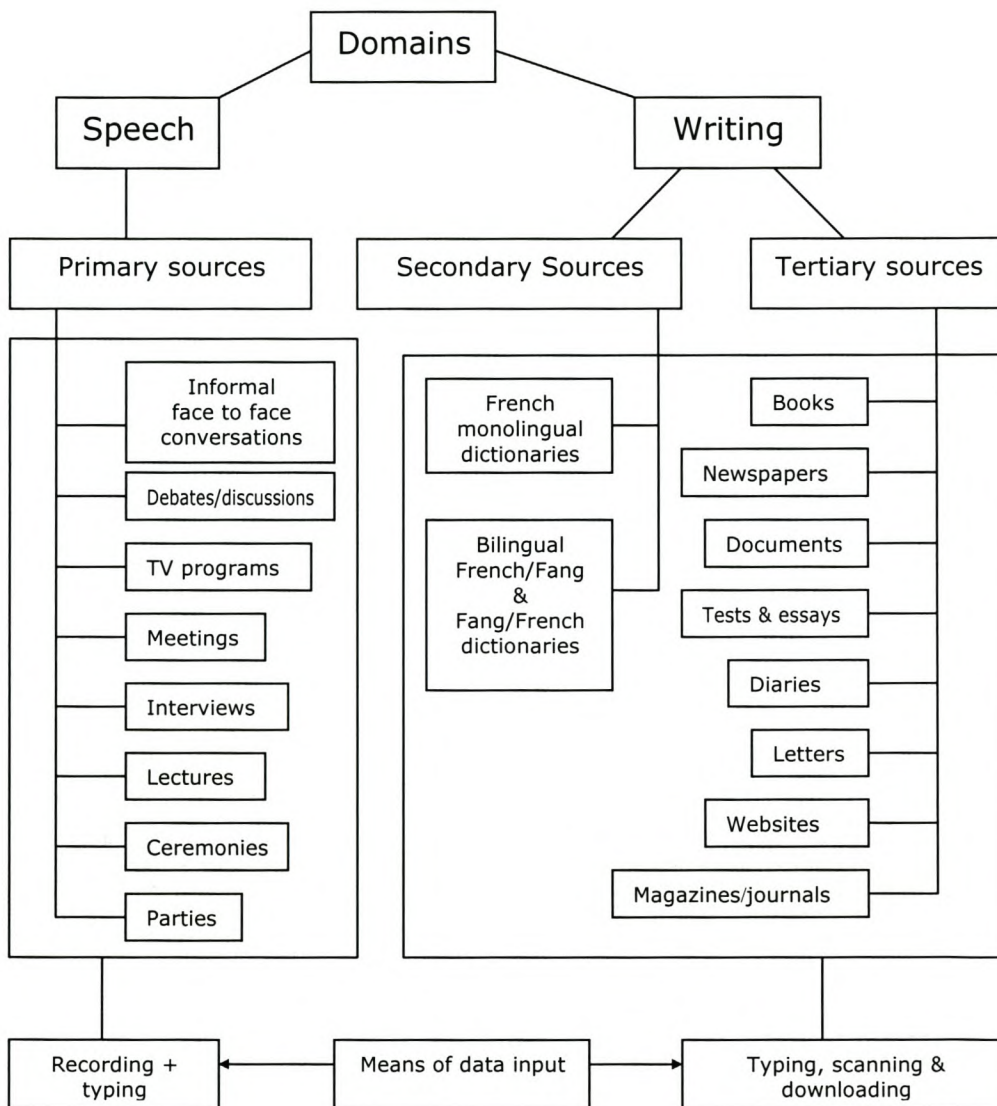
All these spoken texts categories should be used by the lexicographer if possible. Atkins, Clear and Ostler (1992:2) add that the printed monograph or a work of narrative fiction is the model for the notion of a "text". It manifests several apparently characteristic criteria:

- It is discursive and typically at least several pages long.
- It is integral.
- It is the conscious product of a unified authorial effort.
- It is stylistically homogeneous.

Although these types of sources are not often found in the Gabonese context, some available written texts from the target users could be useful in this regard. It is for instance the case of *Les matitis*, a literary book from Ndong Mbeng (1997), a young Fang speaker which narrates the daily live of some boys in a poor area of Libreville. In a style completely free and informal, Ndong Mbeng uses in this book a French very close to the way it is used by most young Gabonese people nowadays.

The whole collection process, from the primary to the tertiary sources, can be represented by the following schema.

Schema of data collection for a French/Fang standard bilingual dictionary



3. 3. 3. Material preparation

3.3.3.1. Introduction

In the material preparation phase the lexicographer has to prepare the collected material for the next steps of the lexicographical process. Once the material acquisition has been completed, the next important step concerns the transfer of the collected and selected material, written and spoken, from the original medium to a computer-readable form (Gunnel Engwall, 1994:53). In the case of the oral material, the recordings have to be transcribed and scanned into the computer for eventual inclusion in the corpus (Gouws, 2001:69). The recording will then have to be transcribed on a word processor using the available (or created) orthography, and be saved instantly in computer files (Prinsloo, 2000:4). The electronic text must, as far as possible, be identical to the original version. It should include all elements of a specific text sample, even grammatical words, proper names, and numbers; elements that were previously often excluded from collected text material. As they form an integral part of the running text, such elements should of course be included in the encoded text and consequently in the corpus (Gunnel Engwall, 1994:53).

3.3.3.2. Preparation of written material

According to Prinsloo (2000: 4) when it comes to existing written material, there seems to be three ways of entering them into computer files:

-
- Electronic transfer, e.g. downloading of a variety of well-selected documents from the Internet or retrieving texts which already exist on computer disk into the corpus.
 - (Re) keyboarding, i.e. typing of hand written documents or even printed matter into computer files.
 - Scanning of printed matter into computer files by means of the so-called OCR (Optical Character Recognition) process using computer software such as OmniPage.

Amongst these three ways of entering written material into computer files, the second way (Keyboarding) is the one which should be the priority. Once again the objective is not to compile any corpus for any dictionary, but to compile a corpus that represents the linguistic repertoire of young Fang speakers. The biggest part of the written linguistic matter cannot be scanned, because it is hand written (e.g. letters, diaries, etc.). The only way to get these onto computer files is by keyboarding.

The first way (e.g. electronic transfer) could be an additional means. In fact, some electronic transfers could be done by downloading various written texts and data from web sites built and managed by young Gabonese people in general and young Fang speakers in particular, over the World Wide Web (Internet), e.g. *www.afanotsaga.fr.st*, *www.nguema-anicet.fr.fm*, *www.minvoul.net*, *www.ndongs.com*. These web sites, written in French are very good illustrations of the free way in which young Fang people are speaking and writing French and could be good sources for obtaining terms

related to technology or science (e.g. Internet, html language, download, web site) that the target group uses in French and needs to be named in Fang for the corpus.

The third way could be used by scanning some academic works written by young Fang speakers, when these works are typed (with typing machine) or computerized. In my view, academic works should be used carefully. Usually these works are done with the assistance of supervisors and they are written at a level of French which does not really reflect the linguistic competence of the authors (students). For most cases, these works do not really reflect the linguistic needs of the target group. I agree with Prinsloo (2000:11) who said that it is not the hardware (the scanner) itself which is crucial in the process, but the Optical Character Recognition (OCR) software which determines the quality of the end product. In fact, up to the present, the most sophisticated OCR systems still do not recognise all language characters.

3.3.3.3. Text encoding

According to Prinsloo (2000:4) text collection is often (at least for corpora of major European languages) followed by text encoding, i.e. the raw text is supplemented by a series of so-called standard corpus preprocessing annotations. He specifies that text encoding can consist of any combination of word tokenisation, part-of-speech tagging, lemmatization, syntactic parsing and marking up.

- By *tokenisation* he means segmenting a text containing conjunctively written words into free-standing words.

-
- *Part-of-speech tagging* is assigning a word class to all the words in the corpus. With reference to Kennedy (1998:21), Prinsloo argues that the original raw text can also be annotated or pre-processed linguistically to show the word class of each word in the text by means of a grammatical tag or label which is attached to each word. He adds that among others, part-of-speech tagging provides crucial data for lemmatization, parsing, or advanced concordancing.
 - Part-of-speech tagging, together with detailed morphological information, enables the lemmatization of a corpus. According to Hartmann & James (1998:83), the *lemmatization* should be understood as the reduction of the paradigm of variant word forms to a canonical form, e.g. the inflected forms (-s, -ed, -e, -ing, etc.) of English verbs to the infinitive. Hence, a "lemmatizer" (also called "morphological analyzer") merges a certain paradigm of variants into a single canonical form. Lemmatization is especially useful for lexicographic purposes (Prinsloo, 2000:4).
 - If tagging deals with words, parsing deals with entire sentences. A structural analysis of sentences is known as "syntactic parsing". Corpora can also be parsed to show the sentence structure and the function in the sentences of the different word classes (Kennedy, 1998:21). *Syntactic parsing*, according to Prinsloo, constitutes an important stage in, for example, machine-aided translation or dialogue systems.

-
- The four standard pre-processing annotations discussed so far (tokenisation, part-of-speech tagging, lemmatization and syntactic parsing) are used to encode “detailed factual or interpretive data about the original text for the purpose of information retrieval” (James et al. 1994:24 as cited by Prinsloo, 2000:4). Existing electronic texts, however, are often already annotated. Yet such annotation, better known as “markup”, operates on a different level, as here “text features” are encoded. The popular way of encoding various electronic text features such as typeface, line breaks, word breaks, sections, paragraphs, headings, and all other aspect of page layout, is the *Standard Generalised Markup Language* (SGML).

As far as the corpus of the dictionary model I propose is concerned, all steps regarding the text encoding stressed by Prinsloo, will not be applicable and necessary at the same level.

- The tokenisation could be done only for Gabonese words used today as part of Gabonese French.
- The part-of-speech tagging will not be a necessary step, because the tagging already made in existing French dictionaries could be used directly.
- The same situation applies to lemmatization, since the planned dictionaries will not give much place to morphological analysis.
- The syntactic parsing will be useful for understanding the structure of sentences used by the target group.

- The mark-up will allow the lexicographer to ensure a good layout of the corpus and to prepare it for the processing phase.

3.3.2.4. Corpus query tools

As Prinsloo (2000:5) states, corpora are of no use without powerful "corpus query tools". As a minimum requirement, such tools must be able to:

- Deal with huge numbers of text files.
- Handle files stored in plain texts as well as in markup format.
- Calculate basic statistics.
- Present alphabetical and frequency wordlists.
- Provide concordance lines.

Altenberg & Eeg-Olofsson (1990: 7), by citing the three main phases of the project "Phraseology in Spoken English", have stressed some requirements that can be needed for the model I propose:

- Retrieval of recurrent word combinations in the corpus. This query will allow the lexicographer to have a better idea regarding the linguistic habits of the target group.
- Elimination of irrelevant combinations on the basis of structural criteria.
- Production of concordances and frequency lists.
- Typological classification of the resulting combinations with regard to their grammatical structure (word-class, phrase category, clause function) and degree of idiomaticity (complex words, lexicalized phrases, free construction).

-
- Calculation of the collocability of frequent word combinations.
 - Description of the positional, functional and prosodic tendencies of the collocations within the tone unit/utterance, their discourse function in the speech situation, and their distribution in different speech types.

For requirements of the corpus model I propose, the production of concordances and frequency lists from the corpus will allow the lexicographer to find the French words most used by young Fang people. The retrieval of recurrent word combinations, the calculation of the collocability, the description of the positional, functional and prosodic tendencies of the collocations will allow him to have a better idea regarding the linguistic habits of the target group.

These steps will play an important role in the selection of the macrostructure and the treatment of the microstructure as confirmed by Prinsloo (2000:2) who said that “on the macrostructural level corpora provide crucial information for the creation of the lemma-sign list of a dictionary, and on the microstructural level corpora enable lexicographers to enhance the accuracy of the dictionary articles tremendously. In the treatment of articles for instance, definitions, examples and illustrations could be retrieved from the corpus queries. There are quite a number of available software packages that are able to perform these tasks, e.g. *Corpus Bench* from Denmark, *MonoConc* from the US, *WordSmith Tools* from England, or the *Access-based program* developed at the University of Pretoria”

(Prinsloo, 2000:5). Prinsloo (2000:5) insists that a “large *query terminology* is connected with corpus query tools, of which the terms **token** or **running words**, **types**, **hapax legomena** and **KWIC concordance** are absolutely basic”. Each term corresponds to a specific task query to the computer:

- **Token** or **running words** stand for the *total number of items* in a corpus.
- **Types** stands for the *total number of different items* in a corpus.
- **Hapax legomena** refers to those *items which appear only once* in a corpus.
- **KWIC concordance** is the acronym for “keyword-in-context concordance”, and is a word or phrase extracted from a text and listed in alphabetical, frequency or other order, together with the words occurring in its immediate environment (Hartmann & James, 1998:79 cited by Prinsloo 2000:5).

“Linguistic corpora seek to further our understanding of language through the analysis of the quantities of naturally occurring data” (Atkins, Beth Levin & Zampolli, 1994:20). The fruits of this work are of immense interest to all researchers concerned with the study of language (linguists, lexicographers, etc.). That is why, in my view, it is important for any modern-day lexicographer to know how to use computers and query corpus tasks to software. He/she can then manipulate his/her corpus according to his/her needs and objectives, at any time. That is why I think, in most modern metalexicographical works, an important place should be given to discussions regarding

the role and use of computer software in the compilation and the analysis of the corpus.

3.3.2.5. Comments

Gunnel Engwall (1994:54) made some remarks as far as the material collection is concerned. He pointed out six important features that must be considered during the selection process of the linguistic material in corpus building; *form, content, continuity, preparedness, availability and delimitation* of selected texts. I consider that three of those features could play an important role for the compilation of a good corpus for the model I propose, particularly for the compilation of a standard dictionary. In fact, *form, availability and delimitation* of collected data are important for the compilation of any corpus, but in the Fang situation a particular attention should be paid to *content, continuity and preparedness* in order to allow the lexicographer to have data containing researched information, reflecting homogeneous discourse and presented in an efficient way:

- The *content*. A clear distinction must be made between literary and scholarly text. Literary texts are generally characterized by their fictional and scholarly by their factual content.
- The *continuity* should also be one of the researched features. Texts appearing at regular intervals, e.g. daily newspapers, weekly magazines, or television programs could be good sources of data in this regard.

- The *preparedness* is a criterion that could also attract the lexicographer. A prepared text² is one which has been edited, perhaps extensively, before its release.

Another significant aspect identified by Gunnel Engwall (1994:53) concerns some codes and/or symbols such as the marking of italics, diacritics, etc that may affect the linguistic or literary analysis,. I agree that they should be kept apart, as they may be ambiguous and thus confusing. To be included in the corpus they have to be presented in a form compatible with the computer system used. The same applies to most African languages. In the Fang situation, the lack of a specific and unique writing system for Gabonese languages make the situation more complicated. To render these languages as compatible as possible, the spelling system used should be as simple as possible. However, the matter regarding the writing of Gabonese languages, in general, and Fang, in particular, will be discussed in chapter 5, which deals with standardization process and writing system.

3.3.2.6. Corpus updating

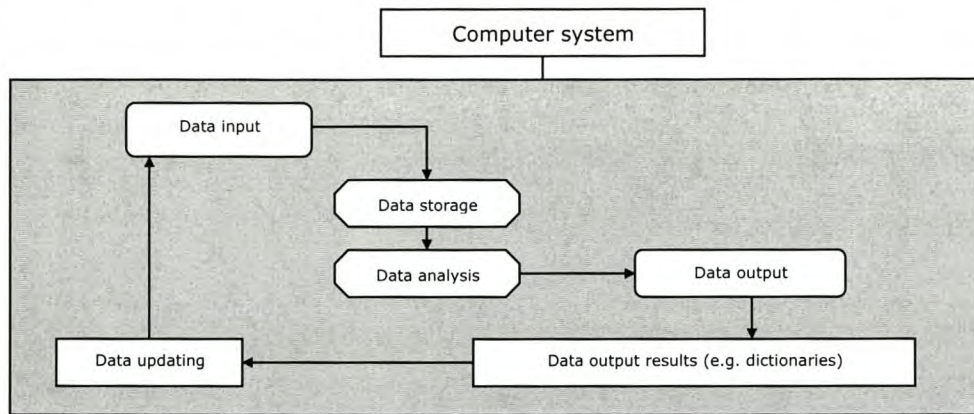
The compilation and analysis of the data are not the end of the process as far as the corpus is concerned. Indeed, the lexicographer must take into account the fact that the compiled corpus will serve for the compilation of other further dictionaries. For that reason the corpus should always be updated and maintained according to the evolution of the language. New words must be added, when obsolete

² The *prepared* text is opposed to a *spontaneous* text which is one produced without the correction process, as applied to a prepared one.

ones will be marked-up in order to allow the lexicographer to be aware of their inclusion or not in the dictionary. Updating the corpus also concerns the way words are written. In fact, if the orthography of some words changes, the lexicographer must also update them in the system by introducing them with the new writing style, but without deleting the obsolete forms. These obsolete forms must be marked-up so that they can be easily identified by the dictionary compiler. When obsolete forms can clearly be retrieved from the corpus they can be used in diachronic dictionaries or in the treatment of diachronic forms in any dictionary. Thus, the inclusion of old dictionaries (lemmas and their entire treatments) in the corpus will play a very important role. Indeed, these old dictionaries will not only allow the compiler to have access to old morphologic forms of words (if they have changed), but also access to their old meanings and pronunciations (if the phonetic transcription is presented in old dictionaries). Therefore, it will be easier for the dictionary compiler to see the evolution of the language thanks to the information available in the corpus.

“In order to overcome the many difficulties involved in language barriers, more attention should be given to the enormous task of the terminologist who supplies multilingual term equivalents for various new concepts in different subject areas and domains” (Alberts, 2002: 91). The updating of the corpus should indeed take into account the technological and scientific environment with the objective of having a corpus as close as possible to the reality of present day language use.

The following schema represents the whole process regarding the updating of the corpus.



3.3.2.7. Corpora and standard dictionaries

The importance of corpora to language study is associated with the importance of empirical data. Empirical data enable researchers to make objective statements, or have a cognitive perception of language. Empirical data also allow researchers to study language varieties such as dialects or earlier periods in a language for which it is not possible to carry out a rationalist approach. Empirical data had been used in lexicography long before the method of corpus linguistics was invented. Samuel Johnson, for example, illustrated his dictionary with examples from literature, and in the 19th century the Oxford Dictionary used citation slips to study and illustrate word usage. Corpora, however, have changed the way in which lexicographers look at language. One of the first advantages lexicographers have when using corpora is the fact that their access to a collection of machine-readable texts allows them to compare texts from different areas. They can easily distinguish items that belong to the standard dialect from

those that belong to the other dialects. As a result, standard dictionaries can be produced and revised much quicker than before, and can thus provide up-to-date information about the standard language. In the same vein, definitions can better reflect the standard language since a larger number of natural examples from the standard dialect can be examined. Instances extracted from the standard corpus can be easily organised into more meaningful groups for analysis, e.g. by sorting the right-hand context of the word alphabetically, it is possible to see all instances of a particular collocate together. Furthermore, because corpus data contain a rich amount of textual information (e.g. regional variety) it is easier to tie down usages of particular words or phrases as being typical of particular regional varieties.

Nowadays, corpora have a great role in the compilation of standard dictionaries as they enable lexicographers to rapidly produce reliable frequency counts and to subdivide these areas across various dimensions according to the varieties of language in which a word is used. By doing so the lexicographer is able to decide if an item has to be included or not in the standard dictionary. Finally, the ability to call up word combinations rather than individual words and the existence of mutual information tools which establish relationships between co-occurring words means that researchers can treat phrases and collocations more systematically than it was previously thought possible. A phraseological unit may constitute a piece of technical terminology or an idiom, and collocations are important clues to specific word senses.

3.3.2.8. Concluding remarks

The compilation of a dictionary is a difficult and time-consuming task and needs a sense of balance. To reach this objective as far as possible, lexicographers, nowadays, use computer facilities that assist them in various lexicographical tasks, including the preparation of lexicographical evidence from many sources, the recording in corpus of the relevant linguistic data, the editing of lexicographical entries, and the dissemination of lexicographical products (Zampolli, 1994:4). This huge participation of data processing in the lexicographical process shows that it is essential, even vital for any modern lexicographer to have basic knowledge of the computational input of data in the corpus, the manipulation or treatment of data by software and the maintenance or updating of the corpus. Besides computer facilities, it is necessary to acknowledge that, in modern lexicography, corpora play a very important role in the successful compilation of dictionaries. For these reasons, nowadays, the phase of collection and analysis of the data is an essential step in the dictionary compilation process.

The main goal of this chapter was to review some steps in the compilation of a corpus. These steps are guided by an essential fact: the linguistic needs of the target group or potential users of the planned dictionary. Indeed, the material collected and the sources of this material must always have a strong link with the target users.

Nowadays the term "corpus" nearly always implies the additional feature "machine-readable". This was not always the case as in the past the word "corpus" was only used in reference to printed

(sometimes written) texts. Today few corpora are available in non-machine-readable form. Machine-readable corpora have the advantage over written or spoken formats in the way that they can be searched and manipulated at speed, and they can easily be enriched with extra data.

There often is a *tacit* understanding that a corpus constitutes a standard reference for the language variety that it represents. This presupposes that it will be widely available to other researchers, which is indeed the case for many corpora. One advantage of a widely available corpus is that it provides a yardstick by which successive studies can be measured. So long as the methodology is made clear, new results on related topics can be directly compared with already published results without the need for re-computation. A standard corpus also means that a continuous database is being used. This implies that any variation between studies is less likely to be attributed to differences in the data and more to the adequacy of the assumptions and methodology contained in the study.

The conversion of written or spoken texts into machine-readable form is not a simple task. Even with a basic machine-readable text, issues of encoding are vital. In tone languages such as Fang, the issue of tones remains a problem. If the matter of accent has been solved for languages with accents like French, thanks to IBM-compatible computers that are capable of handling accented characters, the notation of tones remains a problem for the compilation of corpora in tone languages. Therefore, for maximum efficiency, either the tonal characters should also be encoded in computer keyboards, or the

matter can be solved by omitting tones in the orthography. That strategy was used by corpus compilers for accent languages. My position in this particular matter will be given in the chapter dealing with the alphabet and standardization.

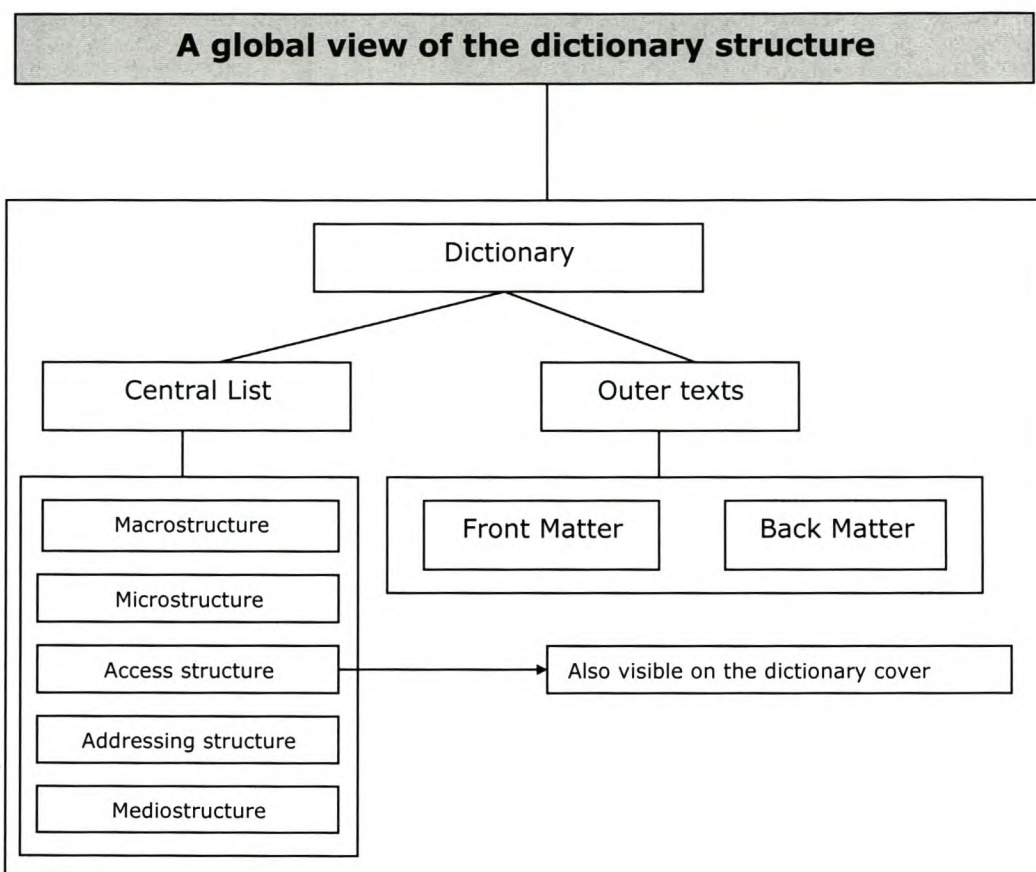
As far as the compilation of a corpus for a French/Fang dictionary is concerned, the first part of the operating system will concern data in French. From that corpus, data will be retrieved for the treatment of the first monoscopal part of the proposed dictionary (French/Fang). A further second monoscopal part (Fang/French) should also be planned when compiling a corpus. For that reason the collection and the updating of the corpus should also be done in Fang in order to make the compilation of a further monoscopal Fang/French dictionary easier. Since the planned dictionary projects to give an important treatment to Fang equivalent items, the Fang corpus will help in the lexicographical treatment of Fang translation equivalents. Both corpora (French and Fang) could also be used in the future for the compilation of monolingual dictionaries (in French or in Fang).

CHAPTER 4: Aspects of the dictionary structure

The common belief that the making of a dictionary starts with defining words is as naïve as the idea that the erection of a building starts with the purchase of the construction materials. If one can imagine ordering tons of steel girders, cement, bricks, windows glass, plumbing fixtures, and electrical wiring without having any plan of the building for which these materials are to be acquired, one will see the absurdity of the notion that dictionaries begin with defining. (Landau, 1984:226)

4.1 Introduction

Dictionaries as any organized system have a structure, which is represented by its component parts in terms of its overall design and contents of individual entries. To evaluate and describe the structural component of dictionaries, terms like **macrostructure** and **microstructure** have been developed in metalexigraphy. This chapter has the purpose to present some aspects of the dictionary structure through the **data distribution structure** in the **central list** and the **outer texts** of the dictionary. This presentation will be complemented by an account of some other aspects of the dictionary structure such as the **access structure** and the **mediostructure**. The presentation of the dictionary structure in this chapter is based on the general theory stressed by Hausmann and Wiegand (1989). It will enable me to better understand Hausmann and Wiegand's approach in order to find the right way this theory can be applied in my proposed dictionary model.



4.2. Data distribution structure

4.2.1. General remarks

At the phase of the dictionary conceptualization, before starting the writing of the dictionary, the lexicographer has to know or decide in which part of the dictionary each available and useful type of data should be set. Except for the material acquisition phase and the material preparation phase discussed in the chapter regarding the database, the data distribution program is one of the main phases in the dictionary conceptualization plan and can be considered as one

step of the material processing phase. The data distribution structure of a dictionary determines the specific position for each data type in the dictionary as so-called *carrier of text types*. Some data will be included in the texts accommodated in the **front** and **back matter** while other data will be included in the articles, the texts constituting the **central list** of a dictionary (Gouws, 2001b: 70). The data distribution structure also allows the lexicographer to make provision with regard to the different search zones within the dictionary articles. As such, it determines the article internal presentation and the diverse search zones within the dictionary articles. For instance, if some data categories (e.g. pronunciation, paraphrase of meaning, translation equivalent, etc.) have to be included the lexicographer has to make provision for article slots or search zones in the dictionary article for these data categories. This, to a certain extent, corresponds to Wiegand's (1996d) concept of **micro-architecture**. The central list of a dictionary is no longer the only venue for the presentation of data or for the occurrence of lexicographic texts. Although the central list remains an important and compulsory text, data can also be presented in the outer texts, i.e. the texts preceding and the texts following the central list. In this section a discussion will be made concerning the distribution of the data categories in the main parts of the dictionary with an implementation of the model I am proposing. Various specific points of the model I propose will be highlighted. Particular attention will be given to some parts of the **central list**, i.e. the **macrostructure**, the **microstructure**, the **addressing structure**, as well as the **outer texts** (front and back matter) of the dictionary.

4.2.2. Data distribution structure in the central list

4.2.2.1. Introduction

In a textual perspective, the word *dictionary* has at least two meanings: the whole book and the word list, which constitutes the main part of the book. The word list (or central list) has several important units and structures (Hausmann & Wiegand, 1989:328). The whole set of lexical entries constitutes the central list (Dubois & Dubois, 1971: 57). According to Gouws (2001b: 82) the central list of the dictionary consists of article *stretches* and each article stretch includes a variety of articles, which function as texts in their own right. The central list hosts the most salient structural components of the dictionary. A model for a new dictionary and any dictionary specific lexicographic process will necessarily have to ensure that these structural components are established and used in a functional way. Gouws (2001b: 83) adds that research in the field of metalexigraphy has led to the identification of at least five structural components to be negotiated in the central list of a dictionary, i.e. the **macrostructure**, **microstructure**, the **access structure**, the **addressing structure** and the **mediostructure**. As far as the relationship between data distribution structure and central list is concerned, only aspects related to the macrostructure, the microstructure and the addressing structure will be discussed. The access structure and the mediostructure, which are not parts of the data distribution structure, will be discussed in different sections.

4.2.2.2. Macrostructure

4.2.2.2.1. General remarks

The basic purpose of a bilingual dictionary is to coordinate the lexical units of one language with those lexical units of another language, which are equivalent in their lexical meaning. The first language, with which lexical units from the other language are coordinated, is called the source language; the order of lemmata in a bilingual dictionary is given by the source language (Zgusta, 1971:298). All the lemmata of a dictionary form the macrostructure. If the dictionary has only one word list, the alphabetic access structure and the macrostructure are identical. If the dictionary has additional word lists, the central list and the additional word lists are all partial structures of the macrostructure (Hausmann & Wiegand, 1989:328). The macrostructure is the collection of lemmata included as part of the central list. Although the macrostructure usually contains a list of words, it may not be defined as a collection of words (Gouws, 2000:22). The macrostructure is the selection of lexical items to be included in the dictionary as lemma signs. They become the primary treatment units of the dictionary (Gouws, 2001b: 83). The type of the dictionary, above all, its intention and purpose should govern the selection of the prospective lexical items, which will be included in the bilingual dictionary. In the case of a dictionary, which intends to help the user to understand texts couched in a foreign language, it will be clear that the occurrence of the lexical units in those texts is the first factor which determines the selection of the lemmata (Zgusta, 1971:309). The macrostructure is the only structure, which may contain all the units to be treated. It is

the macrostructure that determines under which lemma the lexicographical item is to be found (Hausmann and Wiegand, 1989:336). In this section by macrostructure I mean the collection of lexical items selected as lemmata in a dictionary.

4.2.2.2.2. Macrostructure in Fang dictionaries

4.2.2.2.2.1. Introduction

Before I dissert about a new macrostructural model for any future Fang dictionary, a brief overview of the macrostructure of some existing Fang lexicographical works will be made. This synopsis will allow me to suggest which aspects of these dictionaries should be improved or adapted according to the model I propose. The overview mainly focuses on three publications: the *Encyclopédie Pahouine* of Largeau (1901), the *Lexique fān-français* of Martrou (1924) and the *Dictionnaire fang-français et français fang* of Galley (1964). These publications are, up to the present, the most important lexicographical works compiled in Fang.

4.2.2.2.2.2. Macrostructure in the *Encyclopédie Pahouine* of Largeau

Compiled by the French colonial administrator V. Largeau, the "Pahouin Encyclopedia" is an important work of more than 4900 articles. It is a strict alphabetically arranged dictionary that is, in the first place, for French speaking people who want to translate French into Fang. In fact, according to the compiler, this dictionary was compiled to help colonial administrations that were no longer disposed

to place faith in interpreters¹. Largeau's work has a monoscopical character. French is the source language of the dictionary and lemmata are presented in that language. Lemmata written in small capital letters (only the initial is a normal capital letter) are presented according to the lemmatization principles established in the European tradition, e.g. verbs are lemmatized under their infinitive form or under the first person singular of the present indicative, whereas nouns are entered under their complete form (Mavoungou, 2001a: 124). Different types of lemmata from different fields (e.g. ethnology, religion, oral literature, metallurgy, astrology, etc.) and multiword lexical units (e.g. *tout à coup*, *tout à l'heure*, *arc-en-ciel*, etc.) are also included. Several of these lexical units do not reflect the everyday conversations of Fang people, but the linguistic habits of the target users (i.e. French colonial administrators).

4.2.2.2.3. Macrostructure in the *Lexique Fān-Français* of Martrou

The Fang-French lexicon compiled by Archbishop L. Martrou is a lexicographical work of more than 3400 articles, with Fang as source language. Lemmata are arranged in strict alphabetical order and written in bold and small characters and only the initial is in capital letter (e.g. **Asara**, **Asas**, **Asè**, etc.). In the front matter texts the compiler specifies that the orthographical system applied to the transcription of lemmata in Fang have been established according to the principles expounded by Ch. Sacleux in *Essai de Phonétique*. With

¹ Because of many mistakes made by interpreters during their translation there were many troubles between indigenous people and colonial administrators.

this spelling system he tried to write out the phenomenon of tones in Fang by using some diacritic signs (e.g. **Efõn**, **Ekènœ**, **Ekœ'roeboe**, etc.). Different categories of lemmata are included in the macrostructure (e.g. nouns, verbs, adjectives, pronouns, etc.), but most of the words selected as lemmata come from fishing and hunting fields. This could be explained by the fact that the Fang people (i.e. Fang-Atsi) from whom the data was collected exert their main activities in these fields.

4.2.2.2.4. Macrostructure in the *Dictionnaire fang-français et français-fang* of Galley

Also known under the name of "The Galley", the dictionary Fang-French and French-Fang of Samuel Galley can be branded as the most important lexicographical work compiled in Fang up to the present day. This biscopal dictionary of more than 13900 articles has a double macrostructure: the first macrostructure of the dictionary is in Fang (i.e. Fang-French section) and the second one (i.e. French-Fang section) in French. The macrostructure is presented in strict alphabetic order and lemmata are entirely in bold and capital letters (e.g. **BANA**, **BANÉ**, **BA-NYE**, etc.). Galley did not specify the orthography system he used for the transcription of lemmata in Fang, he nevertheless tried to take into account the phenomenon of tones in this language. After each lemma-sign in Fang follows, in parentheses, the indication of the melodic curve or tone²: e.g. **BARA** (b), **BEBÉE** (h), **BEBEBÉ** (m), etc. Only three punctual tones are identified (i.e. low, medium and high).

² (b)= low tone; (m)= medium tone; (h)= high tone

As far as the selected lemmata are concerned, the compiler has included words used in daily language usage such as geographic names, plant and animal names, proper names and expressions.

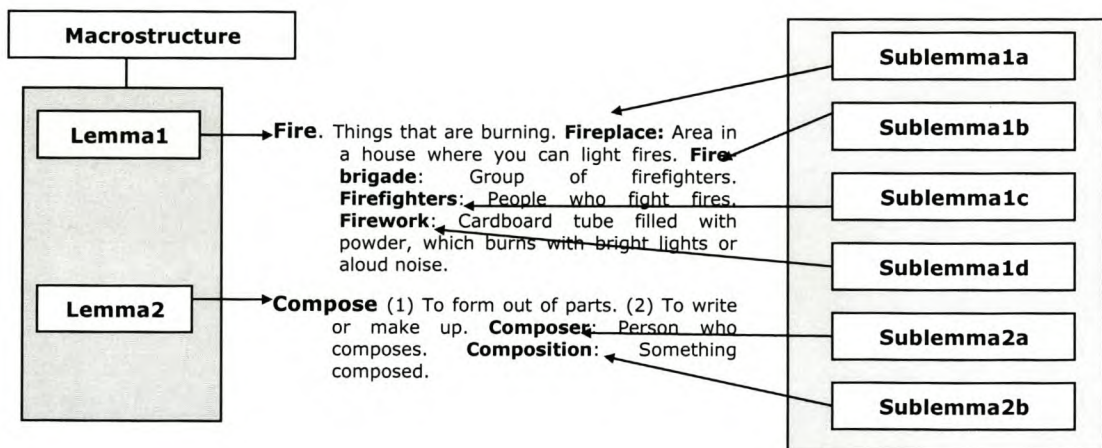
4.2.2.2.3. Different types of macrostructures

The successful retrieval of information in a dictionary often depends on an unimpeded access to the lemma-sign. Lexicographic planning includes a clear-cut decision regarding the type of macrostructure to be presented in the dictionary. The arrangement of lemmata is of primary importance. The typological characteristic of general dictionaries determines that they should display an alphabetical arrangement. However, lexicographers have to make a distinction between a straight alphabetical macrostructure and a macrostructure with a sinuous lemma file. Where a sinuous lemma file is presented a further distinction has to be made between **niching** and **nesting** dictionaries (Gouws, 2001b: 85). *Niching* dictionaries have a *strict-alphabetical clustering* of lemmata or articles that may or may not be semantically related. Nesting dictionaries have a clustering or listing of lemmata or articles, which stretches the rules of strict-alphabetical ordering in order to exhibit morphosemantic relations between words (Hausmann and Wiegand, 1989:336). Hausmann and Wiegand (1989:336) confirm this by saying that the macrostructure may be presented differently within the two dimensions that are available so that we can distinguish different designs. The common feature of all these designs is the orientation from top to bottom, which is conditioned by the writing system. In *straight-alphabetical* dictionaries,

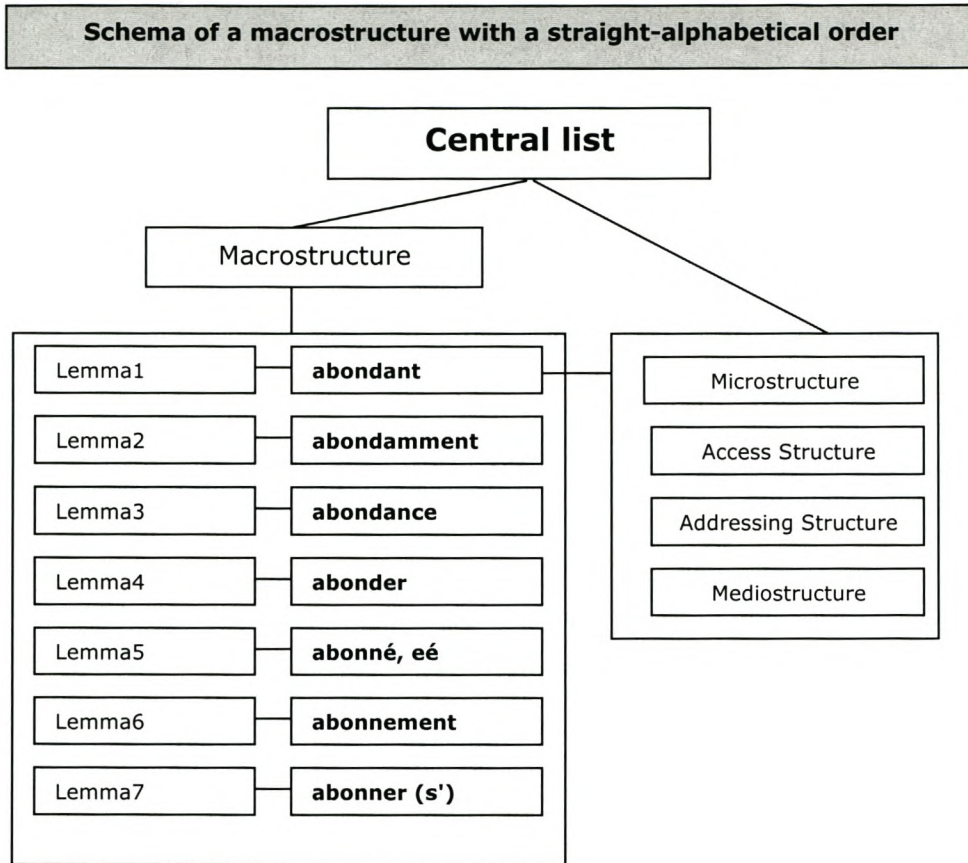
this orientation is strong and the lemma file is arranged vertically. Elsewhere we find sinuous lemma files. This is the case in *niching* and *nesting* dictionaries. It is important to emphasize that niching and nesting ordering concerns sublemmata. In most dictionaries, these sublemmata are also given in bold.

One of the ways a lexicographer can attain economy of space in a particular dictionary article is to reduce the design by the utilization of textual condensation procedures of the data presentation. The presentation of specific lexical items as sublemmata is one way to make use of an important space-saving tool.

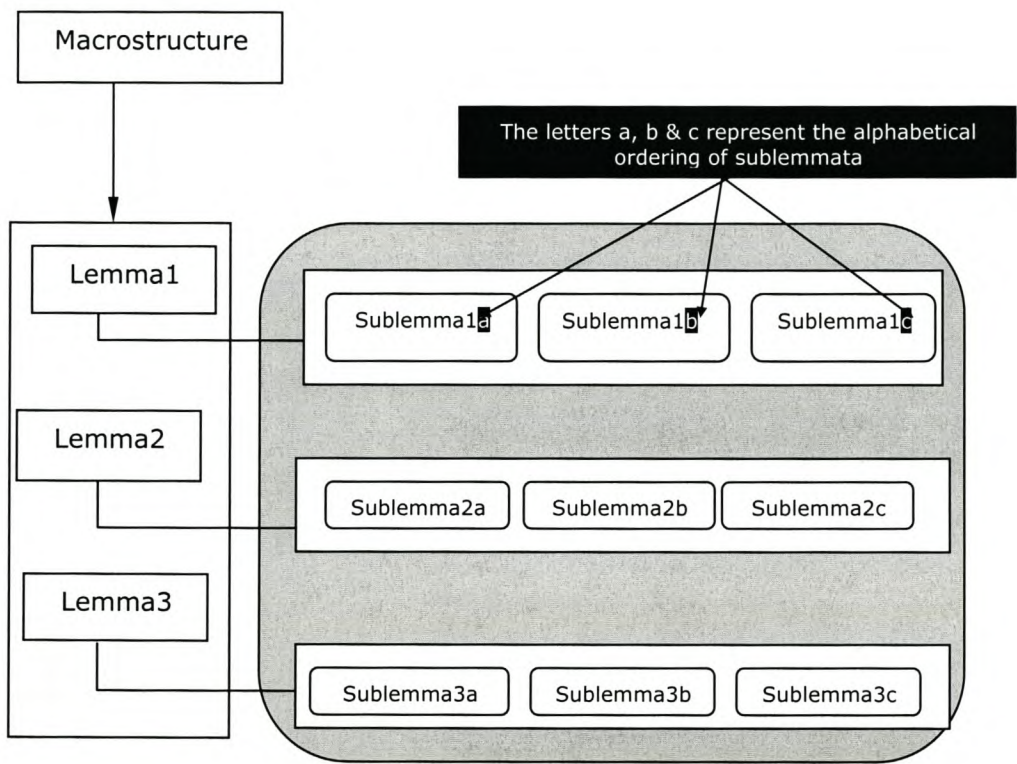
Illustration:



Another aspect of the macrostructure that must be highlighted is the fact that the alphabetical arrangement can follow one of two alternative principles: word-by-word, or letter-by-letter. The difference can be seen when there are lemmata comprising more than one graphical word (Gouws, 2000:28). Svensén (1993:223) explains this by saying that the letter-by-letter method treats lemmata as if they were written continuously, while the word-by-word method takes account of the spaces, which are always ranked before the letters. Thesauri and other dictionaries with a thematic approach have a non-alphabetical ordering of the macrostructure. However, most dictionaries display an alphabetical ordering of their lemmata. This includes standard, comprehensive and learner's dictionaries. Hausmann and Wiegand (1989:336) add that a sinuous lemma file also proceeds roughly from top to bottom. But this orientation is weaker, because it is often disrupted when two lemmata follow one another on the same line or where there is lemma-clustering. The lemma-cluster results from space saving textual procedures.

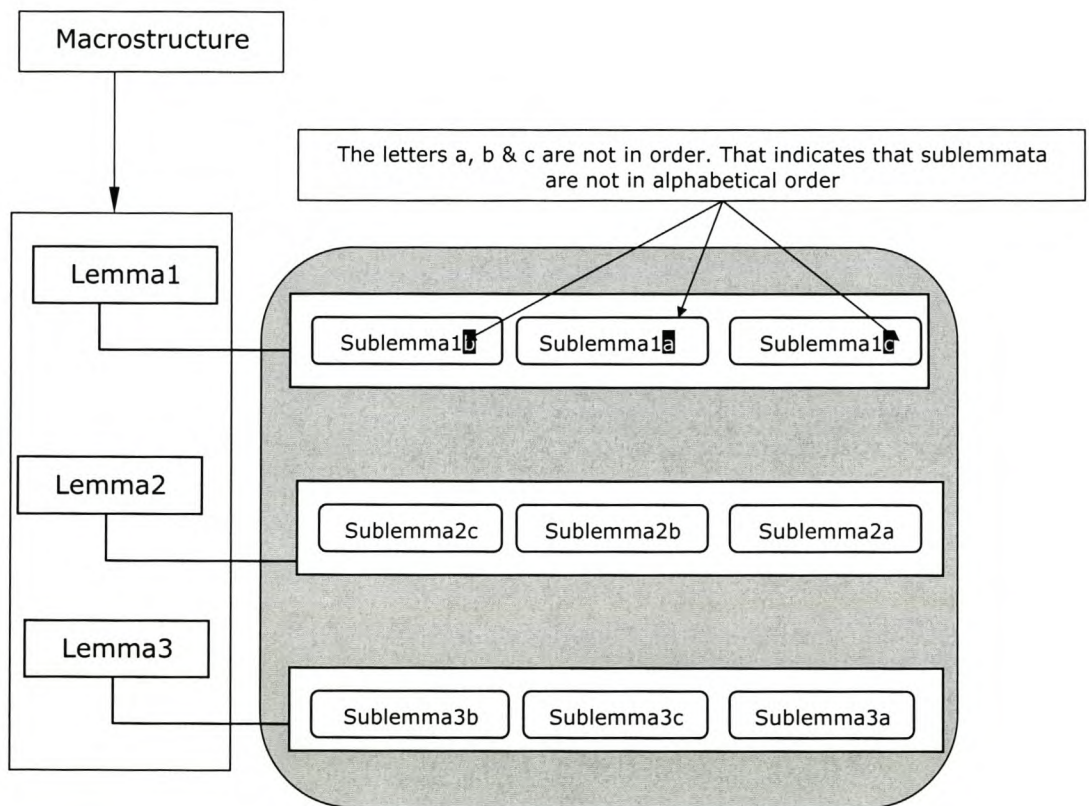


Schema of a macrostructure with a *Niching* ordering



In the niching ordering, sublemmata are arranged in alphabetical order; however, these sublemmata may be or may not be semantically related.

Schema of a macrostructure with a *Nesting* ordering



In the nesting ordering, sublemmata are arranged in both alphabetical and non-alphabetical order; however, these sublemmata are arranged according to their morphosemantic proximity.

4.2.2.2.4. Proposed macrostructure

For the theoretical model I propose the macrostructure should be presented in a strict alphabetical order, the so-called word-by-word method of Svensén. The strict alphabetical ordering implies that lemmata display a vertical macrostructural arrangement and all these lemmata are positioned alphabetically. This presentation has considerable advantages:

- The users can exploit their alphabetical knowledge to quickly find the needed lemma sign.
- The alphabetical order puts the lemmata in a conventional system of article stretches from the letter A to Z. Lemmata are in a strict-initial alphabetical ordering. In such a classification, one can verify the presence or the absence of a term without too much difficulty, once its orthography is known. On the other hand, conceptual dictionaries, which organize the macrostructure according to a semantic order (e.g. areas of interest, domains, etc.), do not ensure the entire treatment of the terms and cannot be consulted easily without the help of an index (i.e. a return to the alphabetical order). Therefore, the alphabetical order stays the dominant and perhaps the exclusive way. The alphabetical order remains the most commode and rigorous; that presentation is an insurance of exhaustiveness. It easily allows the verification of the presence of a form or a word in the dictionary (Mitterand, 1963: 114).

-
- Searching is greatly facilitated if the lemmata are always to be sought on the extreme left of the column, and their appearance is not impaired by divided lemmata (Svensén, 1993:224).

This method, which demands lesser dictionary user skills, also saves the lexicographer work. He does not need to decide which words can or should be included in the nesting or niching presentation, and the material is easier to handle during the editing, coding, sorting, layout and revision (Svensén, 1993:225). Therefore, the macrostructure of the proposed dictionary will not have any sublemma. Any item that could be presented as sublemma, in niching and/or nesting dictionaries, will automatically constitute a lemma in its own right in the proposed dictionary model.

In some dictionaries, the data categories presented in the treatment of sublemmata are not always the same as those retained for lemmata. In some situations, the user can wrongly assume that the data categories included for a lemma are valuable for sublemmata in an article stretch. For instance, although part of speech is given for the lemmata in almost all dictionaries, many of them do not give part of speech for sublemmata. I assume that this is also a part of the lexicographer's strategy to save space. That orientation can have negative consequences for the user, who can think that the data categories presented in the treatment of lemmata are also valid for sublemmata. For instance, if a given lemma is indicated to be an adjective, and the part of speech is not indicated for sublemmata in the same article stretch, the user may conclude that these

sublemmata are also adjectives. That is why I propose not to have sublemmata in the proposed article stretches; in order to give a full treatment of each lexical item selected to be included as lemma.

As the proposed dictionary model aims to standardize Fang, lemmata will be presented in the proposed standard dialect (i.e. Fang-Ntumu) and written in the standard alphabet.

4.2.2.2.5. Selection of lemmata

Once the dictionary conceptualization is done and criteria for the selection of lemmata have been established, the lexicographic work can continue with the collection and the selection of lemmata to be included as macrostructural elements. The macrostructure represents the lexicon and this selection may not be made on a random basis. The lemmata have to be drawn from a representative corpus of the specific language (Gouws, 2001b: 85). The inclusion of a specific lexical item as lemma is not motivated arbitrarily but has to be seen as the application of fixed lexicographic criteria. The typological nature of a dictionary also determines the selection of lexical items to be included as part of the macrostructure. Other considerations such as usage frequency³, the register of a lexical item and its morphological nature also play a role (Gouws, 2000:24). The following principles stressed by Zgusta (1971:309) for the compilation of the macrostructure of a bilingual dictionary are in this regard excellent

³ Important studies, by D.J. Prinsloo, Department of African Languages at the University of Pretoria (South Africa), have shown that by choosing the most frequently used words of the language, the chances to include the right lemmata in a dictionary are between 85 & 95% (depending on the language).

criteria that can be applied in the proposed dictionary model as follows:

- If the planned dictionary is not to be a big one, it is possible to omit the less known or less used synonyms of the source language if the better known and more used are included. However, in the dictionary model I am proposing this aspect cannot be applied, as the objective is to place at the target users' disposal a dictionary that is as exhaustive as possible. In fact, although several words are less used in French by young Fang-speakers, many of these words should be included in the macrostructure, as long as they are related to fields or domains taught in schools or with aspects from which the target group will need equivalents in Fang (e.g. mathematics, computer, Internet, biology, chemical, chemistry, poetry, philosophy, natural sciences, broadcasting, satellite, cell phone, technology, sport, cinema or movies, etc.). For the model I propose the lexicographer has to make sure that the maximum number of terms from such fields is included in the macrostructure.
- Because the model I am proposing is a standard dictionary, the inclusion of slang expressions or even vulgarities and similar levels of language should be done with caution in such a dictionary. It is strongly recommended to label them very cautiously in order to warn the user. For the model I am proposing, all types of words and expressions must be included, even slang expressions, vulgarities and similar levels of

language, as the target user needs to master the language at all its levels of expression. The most important thing is that the lexicographer labels, with precision and detail, all words and expressions from the mentioned categories. The corpus I am compiling in Fang shows that young Fang people are using, in their every day conversations, many argotic and vulgar terms in French. Compiling a dictionary with them as target users but without including their most frequently used terms will not make any sense. If they use these terms in French, they will also need to use their equivalents in Fang. By including them, the lexicographer will take into account the linguistic needs of the target group in Fang. Finally, the inclusion of this kind of language levels will give the lexicographer the opportunity to make the users aware of the bad use and the negativity of such words and expressions in certain circumstances.

- If the target language of a dictionary of this type is spoken in a different cultural society and in a geographical and/or other extralinguistic milieu vastly different from that of the source language, it will be necessary to consider the target language when the entry words (lemmata) of the source language are selected. Different social institutions, different plants and animals may be unimportant or non-existent in the milieu of the source language while being very important or frequent in the milieu of the target language. That point is applicable where the cultural and geographical gaps between Fang (i.e. target language) and

French (i.e. source language) occur. Indeed, Fang is spoken in a society (i.e. Gabon-Africa) with a different culture and in a milieu largely different from that of French (i.e. France-Europe). Many important realities (e.g. snow, horse, castle, etc.) known in the French society will not be relevant in the Gabonese context, in general, and the Fang one in particular. In my opinion, as the proposed model is a standard dictionary, these sort of terms do not have to be systematically excluded. However, the lexicographer should make sure that their inclusion in the macrostructure will not be done to the detriment of terms more relevant in the culture of the target group. These terms could be useful for academic purposes⁴, and to the general culture of the target users. It should be noted that terms related to unknown realities in Gabonese society a century ago, are nowadays part of the lifestyle of the Gabonese people (e.g. telephone, television, car, train, flight, football, etc.), and should no longer be considered as exclusively belonging to French or Western culture. In this regard, lexical items related to these realities must necessarily be included in the proposed dictionary and translated from French to Fang.

- For the compilation of a bilingual dictionary in which the target language is not yet fully stabilized, the lexicographer could be confronted with the lack of a considerable number of equivalents

⁴ If Fang is used in the future as medium of instruction, students will probably study some French cultural aspects (e.g. poetry, history, etc.) that they will need to name or express in Fang. The inclusion of terms related to French realities (even Western) will be useful in this regard.

in the source language. In this situation, the selection of lemmata from the source language must be done with a greater than usual respect to the target language. According to Zgusta, the lexicographer should avoid as far as possible the selection of such lexical items that may surcharge the dictionary with neologisms of all sorts. In my view, and for the French/Fang bilingual dictionary I am proposing, the selection of French lexical items not yet well established in Fang should be done without restriction. First of all, if the lexicographer is scared to include too many neologisms in such a dictionary, the objective to develop Fang and allow it to become a modern language as English, French, Spanish, German, Afrikaans, Swahili, etc. will not be reached and the linguistic needs of the target group in Fang will really not be satisfied. In fact, most terms related to certain academic subjects (e.g. chemistry, computing, etc.) are not fully established in Gabonese languages. Secondly, there is no doubt that if the actual modern realities have been named in other (mentioned) languages, they can also be named in the Gabonese languages, including Fang. One must not forget that even in Western societies for instance, certain realities were unknown, and therefore not named. They had to adapt themselves and find denominations for these new realities when they become established. Every language, at one moment of its evolution encounters such situations. However, on the other hand it will be naïve to deny the evidence that the designation of new realities in Gabonese languages is not yet automatic or

fluent for the majority of the Gabonese people. Even when terms to designate those realities exist, they are not well known or known enough by common speakers (e.g. target users). It is at this level that the inclusion of new terms will really be relevant and useful. For instance, in order to teach or learn mathematics, computing, chemistry, etc. in Gabonese languages, teachers and students will need to get used to terms utilized in these subjects. The French/Fang bilingual dictionary will consequently allow Fang to adapt itself in the actual universal society and promote it as a modern language. In some situations, these new terms are already being used in Fang, but only orally. Their inclusion in the proposed dictionary will give them a written form in Fang and favour their standardization.

In addition to the principles retained by Zgusta for the selection of lemmata, some words that are usually used together could be entered to reflect their real existence, e.g. expressions composed in the same way as English words like *kind-of*, *sort-of*, *used-to*, etc (Swanson, 1967: 65). In the Fang situation certain compound words should be used as lemmata if the corpus prove that they are really used together regularly in the standard language. Some of the following criteria, stressed by Swanson, can guide the lexicographer in the selection of lemmata in a bilingual dictionary. The following can be applicable in the Fang context:

- **Relative frequency.** The most frequent words (and morphemes) given by the corpus analysis should show that the

target group regularly uses those words in French. Therefore, the logic compels the lexicographer to select them as lemmata and thus offer the opportunity to the target users to have their equivalents in Fang too.

▪ **Function words.** The function that some words fulfil in the acquisition of knowledge in Fang or their importance in academic subjects or in the current environment can be an important criterion for the lexicographer to select them as lemmata.

▪ **Semantic criteria.** Words with several senses or synonyms must be considered with particular attention. Indeed, the dictionary must give the user the opportunity to be as competent as possible in the target language. If a word has many senses or synonyms, the user must be able to know its different senses or to recognize its different synonyms when they are used.

▪ **Cultural items.** One of the objectives of the proposed dictionary model is to include items reflecting the culture of the target users. In this regard, Fang will play a determining role for the selection of some lemmata in French. Lexical items related to cultural activities of Fang people must be included (e.g. initiatory rites, traditional ceremonies, etc.).

Concerning the types of lemmata that should be selected, I entirely share Svensén's (1993) point of view, that it may be useful to remember the following rules:

-
- Pay particular attention to the places in the database where the word and its meaning are discussed.
 - Be very careful where the meaning is unknown and no indication of it is given in the context.
 - Be cautious also regarding any compound word whose second element is highly productive, forming compounds that are fairly casually made and have an obvious meaning.
 - If proper names are not to be included solely on account of pronunciation or inflexion, choose only names, which have changed into common nouns or have undergone metonymy.
 - If a word is quite new and has not yet been adopted into the language. It should be put aside and be easily retrievable, so that it can be included in later editions of the dictionary when it has become more established.

All these aspects should be taken into account during the compilation of a French/Fang standard bilingual dictionary.

4.2.2.2.6. Category of the selected lemmata

Generally, words most commonly used as dictionary lemmata are nouns in nominative singular, adjectives in the positive singular (nominative, masculine), adverbs in the positive, verbs in the active infinitive present (Svensén, 1993:66). For the model I propose all the word categories used as lemmata in French dictionaries (e.g. prepositions) could also be selected, because if these words are considered as lemmata in French dictionaries, they need to be given equivalents in Fang. Their selection is, in this case, necessary.

4.2.2.2.7. Form of lemmata

The study of the macrostructure of many dictionaries shows that lemmata can have several or variant forms (Rey-Debove, 1971:122). As far as the form of lemmata is concerned, the theoretical model I propose will adopt Zgusta's (1971:249) approach. According to him all lemmata of the dictionary should be constructed in as uniform a way as possible. Each lemma should be treated as a compartment of its own⁵, containing all information about the respective lexical unit considered necessary for the purpose of the dictionary. The most important part of the dictionary article is the lemma, which is the indication of each respective lexical unit in its canonical form.

The other indications of the article inform the user about the morphological or syntactic combination as well as the class of which the lemma is a member. These second indications will be discussed in the section related to the microstructure, because they are already part of the treatment of the lemma. In existing dictionaries, lemmata are generally printed in bold or semi-bold. In certain dictionaries, particularly those published in France, it is normal to print the lemmata in capital letters. I agree with Svensén (1993:64) that "it is not a recommended practice, as it is of some importance whether or not a word should be written with an initial capital. The basic rule is therefore to show the lemma as it is normally written within a sentence in continuous text"

⁵ It is important to note that in some cases cross-references from one item to another are sometimes necessary and some lemmas are conflated into nests.

For the model I propose, lemmata will be presented in bold and in small letters, allowing the lexicographer to give evidence of the accents and feminine forms. Only terms identified as international conventional acronyms such as ONU (UN in English), SIDA (AIDS in English), URSS (USSR in English), etc. will be presented in small capital letters. The first letter of proper names will also be in small capital letters.

4.2.2.2.8. Functions of lemma signs

4.2.2.2.8.1. Introduction

Lemma signs are entries from which e.g. spelling information can be retrieved. They also show the use of small or capital initial letters. They can also show the recommended way(s) of *dividing a lemma sign* at the end of a line. In certain cases, *pronunciation* is shown by marking the lemma sign. Last but not least, they function as *key* to the information sought (Svensén, 1993:46). They are guiding elements of the article and primary treatment units.

4.2.2.2.8.2. Data on spelling

Many words can be spelt in more than one way. As we are dealing here with an established language with a long lexicographical tradition, it is rarely difficult to determine which spelling is the preferable one and should thus appear as the entry form. Common spelling variants should be included, but normally only in dictionaries that are wholly or partly intended for passive use; they are mostly superfluous in active dictionaries (Svensén, 1993:65). In the Fang situation, as well as in all the Gabonese languages, no long lexicographical tradition exists. Up to

now all dictionaries made in those languages use different spellings and most of the Gabonese languages are not yet standardized. In the chapter related to the standardization, a spelling that should be used for the compilation of dictionaries in Gabonese languages will be proposed. The preferable spelling, for words that can be spelt in many ways, will be determined by the proposed standard dialect. In other words, items to be included in the proposed dictionary model will be written in the standard alphabet and according to the way they are pronounced in the standard Fang dialect, because the intended dictionary model also has the purpose to standardize Fang spelling.

4.2.2.2.8.3. Lemma as key to the dictionary

A lemma is also the head of the dictionary article. It is via the lemma that the user reaches the data sought. In some cases this data is not found in the treatment of the reached lemma, but often after a cross-reference to another lemma. The discussion concerning the search of information in the dictionary will be done in sections dealing with the access and the addressing structures.

4.2.2.2.9. Linguistic units not to be selected as lemmata

4.2.2.2.9.1. Morphemes

In some dictionaries affixes (i.e. prefixes, suffixes, and infixes) are presented as lemmata. Even if these types of linguistic units have a special role in the functioning of languages, they do not always carry meanings by themselves in some languages. Their inclusion as lemmata could force the lexicographer to undertake a specific morphological study of the language in order to find out which

morphemes individually carry meanings. This could be time-consuming for the lexicographer. A further reason for excluding morphemes as lemmata is the fact that they do not function in the same way in the language pair (i.e. Fang and French). For instance, prefixes in French will not necessarily have equivalents in Fang. The role affixes play in French is not necessarily the same in Fang. To use these kinds of morphemes as lemmata in a bilingual dictionary, a specific morphological study of Fang should be done in order to determine which affixes in Fang are the equivalents of those in French. In my view this aspect could be relevant merely in a Fang monolingual dictionary. The ultimate reason I prefer not to include morphemes as lemmata is because they could perplex users who are more used to deal with these sorts of morphological presentations in French rather than in Fang. They could believe that these morphemes work in the same way in both languages. It should be noted that the motivation to avoid the inclusion of morphemes as lemmata should be explained in the user's guidelines of the proposed dictionary model.

4.2.2.2.10. Concluding remarks

One of the major problems that the lexicographer has to solve when compiling the macrostructure of any dictionary is the choice of lexical items to be included as lemmata in the dictionary. This choice has to be based on objective criteria such as the frequency of the item amongst the target user, the role or the importance of the item in the language, the field of the item, etc. Apart from the choice of the macrostructural elements, the lexicographer also has to find the

efficient way in which the selected lemmata have to be presented in the dictionary. Many dictionaries have a sinuous lemma file, while others are arranged in a straight alphabetical ordering. In this project, I propose the use of a straight alphabetical ordering for the proposed dictionary model as that arrangement demands lesser user's dictionary skills. The proposed dictionary model is planned to be a standard translation dictionary that not only can assist the target group to improve their competence in Fang, but also help the standardization of the language. Therefore, the proposed dictionary model will not include sublemmata in order not only to avoid confusion for the user, but also to allow the lexicographer to give a complete treatment to each macrostructural element. It will help the standardization of Fang if lemmata are retrieved from the standard dialect and presented in the standard writing.

4.2.2.3. Microstructure

4.2.2.3.1. General remarks

The microstructure of a dictionary is the set of ordered data that follows the lemma. This set is a constant structure that applies to a programme and code of data applicable to any lemma. The microstructure fundamentally depends on the macrostructure (Rey-Debove, 1971: 151). In this project, I will mainly use the new concept of microstructures in dictionaries stressed by Hausmann and Wiegand (1989). This new concept mainly focuses on different parts of the treatment of lemmata. The new concept attempts to create a theoretical basis that goes beyond the usual way the microstructure of

dictionaries is described. Hausmann & Wiegand (1989:343) have listed different data categories that can be included within the microstructure. They have distributed data inside the microstructure into twelve main categories:

1. Synchronic identification: spelling, pronunciation, part of speech, flexion, and aspect.
2. Diachronic identification: etymology.
3. Diasystematic labeling.
4. Explanatory data: definition, linguistic description, and encyclopaedic description.
5. Syntagmatic data: construction, collocation, example, and quotation.
6. Paradigmatic data: synonymy, antonymy, analogy, homonymy, paronymy, word formation.
7. Other semantic data.
8. Notes.
9. Pictorial illustrations.
10. Ordering devices.
11. Cross-references, references.
12. Representation symbols.

The main purpose of this chapter is to try applying Hausmann and Wiegand's vision of the microstructure on the dictionary model I propose, with specific applications according to my own vision. In this regard, a broad overview of different parts of the microstructure as seen by Hausmann & Wiegand will be presented. However, before such

an exercise, a survey of the microstructures in existing Fang dictionaries will be highlighted.

4.2.2.3.2. Microstructure in existing Fang dictionaries

4.2.2.3.2.1. Introduction

The presentation of a new theoretical model of the microstructure relevant for future Fang bilingual dictionaries will not make sense if a previous analysis of the microstructure in the existing dictionaries in that language is not made. In that regard, three leading dictionaries will serve as foundation of the analysis: the *Encyclopédie Pahouine* of Largeau, the *lexique Fang-Français* of Martrou and the *Dictionnaire Fang-Français & Français-Fang* of Galley. Mavoungou (2001a) and Mihindou (2001) have already undertaken some analysis of those lexicographical works; that is why I will only focus on aspects that will be pertinent to my theoretical model.

4.2.2.3.2.2. Microstructure of the *Encyclopédie Pahouine* of Largeau

The *Encyclopédie Pahouine* of Largeau does not have a typical organization of the microstructure. Indeed, lemmata are treated in different ways according to their meaning, type or function, and sometimes according to the feeling of the compiler. In general three types of microstructural treatment can be observed in the encyclopaedia:

- The first type of treatment, the most simple and the less frequent, shows the equivalent item of the lemma in Fang and thereafter comes the plural form of that equivalent item in Fang.

In some cases the compiler also gives the synonym(s) of the equivalent item and its plural form. No example or illustration is given.

Example:

Animal. tsit. Pl. betsit.

- The second type of treatment gives the meaning of the lemma, in French, the translation of that meaning in Fang (sometimes in parentheses), the synonym(s) of the lemma in French, and the equivalent item(s) of the synonym(s) in Fang, the plural form of the lemma in French and its equivalent items in Fang. When necessary, the compiler also gives different contexts of use of the lemma and its equivalent by giving some illustrations.

Example:

Anguille. Intermédiaire entre poisson et serpent, un mètre long. Elle a des dents. (ezang kos ya nyo, ayap. Ebèlè mèsong). Syn. Poisson à dent. Emvonye. Pl. Anguilles. Mimvonye.

- The third and last type of treatment, the most frequently used, is also the most complex because it includes aspects already present in the first and second types of treatment and combines them with some new features. This kind of treatment particularly concerns lemmata with more than one sense. Thus, a number (e.g. 1^o, 2^o, 3^o, etc.) introduces each sense in French, followed by the equivalent item in Fang. The plural form of the lemma is also given.

Example:

Ver. *etum* pl. vers. *bitum* 1 . Qui mange des corps en decomposition (*da ve na byëm bi sun*). 2 Jeune enfant de 2 à 5 ans. (*emon abele mimbu bibeny ya mimbu mitan*). Syn. Asticot. *Endekh*. Pl. Asticots. *Bindekh*.

Largeau uses two kinds of fonts in the treatment of lemmata: normal font (not bold or italics) for all data written in French and italics for data written in Fang. The compiler did not specify the part of speech and the gender of lemmata in order to allow the user to know for instance if the lemma is a verb, a noun or an adjective or if it is feminine or masculine. In general, data presented in articles are strongly related to ethnographical and cultural aspects such as lifestyle, value system, beliefs, etc. of Fang people. It can explain the reason why the compiler called this work "encyclopaedia".

4.2.2.3.2.3. Microstructure of the *Lexique fān-français* of Martrou

Contrary to Largeau, Martrou opted for one type of microstructure. The lemma (in Fang) is followed by the explanation or meaning in French. In general, the microstructural programme in the Martrou's lexicon is restricted to a set of data in French regarding the lemmata. The compiler did not really give the equivalent items of lemmata, but rather, their meanings or paraphrases of meaning. The meaning is generally complemented with supporting data related to the lemma. This work can be classified in the category called by Tarp (to be published) a monolingual dictionary with a bilingual dimension. That

means a dictionary that has one object language (i.e. the one in which lemmata are presented) and another description language (i.e. the one in which meaning paraphrases are presented). The object language, in this case, will be Fang and French the description one. In the treatment of a lemma with more than one meaning, Martrou distinguishes different paraphrases of meaning with a hyphen. Examples, in Fang, are also given in most cases, followed by their translation in French. Those examples can easily be distinguished from the rest of the data in the article because they are written in italics. Sometimes, the part of speech is indicated, particularly when the lemma is not well known. Occasionally Martrou also uses cross-references when the given definition can be applied to another lemma.

Example:

Afe, adj. Et pr. Ind. Autre, différent (accord sing. De la Cl. V.) *Akoê afe*, une autre Pierre. *Dzal afe*, un autre village. *A f. afœ*. – drôle, original. *O næ afe!* Tu es drôle! *Afe!* Exclamation : oui encore ! Tiens!

4.2.2.3.2.4. Microstructure of the *Dictionnaire Fang-Français et français-fang* of Galley

If the “Galley” is considered as the most important lexicographical work compiled in Fang so far, it is not only because of the number and the variety of the selected lexical items, but also because of the way in which those lexical items are treated as lemmata. In fact, the compiler has tried to be as exhaustive as possible in the treatment of each lemma.

Amongst all dictionaries compiled in Gabonese languages, in general and Fang in particular, the Fang-French and French-Fang dictionary of Samuel Galley is one of those that tried to take into account the phenomenon of tones⁶. Indeed, a tonal indication, in parentheses, regarding the pronunciation, is given, in most cases, after the lemma: (h)= high tone, (m)= medium tone, (b)= low tone. The compiler also indicates the part of speech, the class, the plural form and even the dialect in which the lemma is mostly used. After these data entries follow the equivalent item of the lemma in Fang, the explanation in French, examples (in Fang), their translations (in French) and in some cases, their synonyms. As Martrou and Largeau, Galley also uses the italic font for all entries written in Fang. Loads of cultural or ethnological data related to Fang people are also given. These latter particularly are more abundant when a lemma in Fang does not have an equivalent item (or the concept to which it refers) in French.

It should be noted that the "Galley" is a biscopal dictionary, and the compiler uses the same microstructural procedure in both sections of the dictionary, i.e. the way lemmata are treated in the Fang/French section is the same as in the French/Fang section.

Examples: (the first example is from the Fang/French section and the second example from the French/Fang section).

Etura (b) n.5, pl. *bitura* (vb ture b). Case provisoire, abri de forêt, car on brise les *mezom* sur le toit (ture). Syn.: *ekukula*, *mba*. Proverbe: *nku a nga nyakh mebi etura*. Le *nku* a dit: Je vais demeurer au village, je ne reviendrais plus ici, et il a fait ses besoins dans l'*etura*. Mais en chemin la pluie l'a surpris et il a dû revenir. Morale: il ne faut pas gêner la maison qu'on quitte on ne sait pas ce qui arrivera. Il ne faut pas jeter le vieux *paane*, avant d'avoir reçu le neuf.

Ami amitié, fidélité, *angom* (m), *mvî* (h), *amvî* (h), *mvie* (h), *ngwa* (b). Echange de femmes par amitié, *amvî* (h). Mon ami, *mie* (h), *e mie wam*, *emwi wam*, *angom dam*. Mes amis, *bemie*. Ton ami, tes amis, *nyubele*, *benyubele*. Son ami, ses amis, *ebele*, *bebele*, Notre ami, *angom da*. Votre ami, *ebele vina*. Leur ami, *a mi*, *amwi* (h). Etre amis, se lier d'amitié, *ngoma* (b), *lar angom*, *yan angom*. Celui qui se lie facilement d'amitié, *ngom-ngoma* (bh). Amitié très proche, *angom ne-baghdaa* (mh).

4.2.2.3.3. Different types of microstructures

4.2.2.3.3.1. General remarks

According to Gouws (2001b: 86) the microstructure is represented by all data categories included in a dictionary article as part of the treatment of the lemma sign as well as by the structure and the presentation of the articles. The plan of a dictionary and the data distribution structure should make provision for the inclusion of more than one type of article to be included in the dictionary. He specifies that dictionary research has resulted in the identification of different types of microstructures and the dictionary specific lexicographic process of each project has to instruct the lexicographers with regard to the type of microstructure to be employed in the dictionary. The decision regarding the type of microstructure coincides with the decision regarding the typological classification of the dictionary. A model for a new project should make provision for the choice between three categories of microstructures, i.e. an **unintegrated**, an **integrated** and a **semi-integrated** microstructure. On the basis of the distinction of partial structures of microstructures, Hausmann and Wiegand (1989: 354) think that four types of microstructures may be distinguished, i.e. **integrated**, **partially or semi-integrated**, **unintegrated** and **rudimentary** microstructures.

In this section, I will present all these partial microstructures, and I will suggest the way some of them can be included in the proposed dictionary model. The role these partial microstructures will play in the proposed mode will also be explained.

4.2.2.3.3.2. On the distinction comment on form and comment on semantics

The lemma functions as guiding element of each dictionary article and the microstructural programme orders the entries included as part of the treatment of the lemma in such a way that they can be divided into two major article components, i.e. the *comment on form* and the *comment on semantics*. Every data entry included in the microstructural programme belongs to one of these components (Gouws, 2001:70):

- The **comment on form** includes the orthographic representation of the lemma, i.e. the lemma sign. Additional spelling guidance is included in the comment on form if the lexical item included as lemma has spelling variants. The most typical other entries accommodated in the comment on form are the data conveying information regarding the morphology and the pronunciation of the lemma as well as certain grammatical features. To ensure a systematic retrieval of the information presented in the comment on form it is essential that a dictionary plan should prescribe a fixed ordering of the data types and subtypes. The way in which the entries are presented in the comment on form should also be determined by the reference skills of the intended target users (Gouws, 2001:71). In the proposed dictionary model, the orthographic representation of lemmata will be given in the standard alphabet, and spelling guidance will indicate to the user how lemmata are pronounced in the standard dialect.

- The **comment on semantics** displays a variety of data types. The nature and extend of the comment on semantics are determined by the type of dictionary. It contains all the entries reflecting various aspects of the meaning of the lemma as well as pragmatic values of the lemma. In a monolingual dictionary the paraphrase of meaning usually is the most salient entry in the comment on semantics whereas the translation equivalents have this function in a bilingual dictionary (Gouws, 2001:72). The comment on semantics will primarily reflect meaning and pragmatic values of lemmata in the standard dialect. However, some aspects of specific lemmata could be given in other dialects where their meaning and values largely differ from the standard dialect. That approach aims to establish a norm in the language by making differences between standard or normative value of words that should be used in official and academics contexts, and those that can be used in common usage.

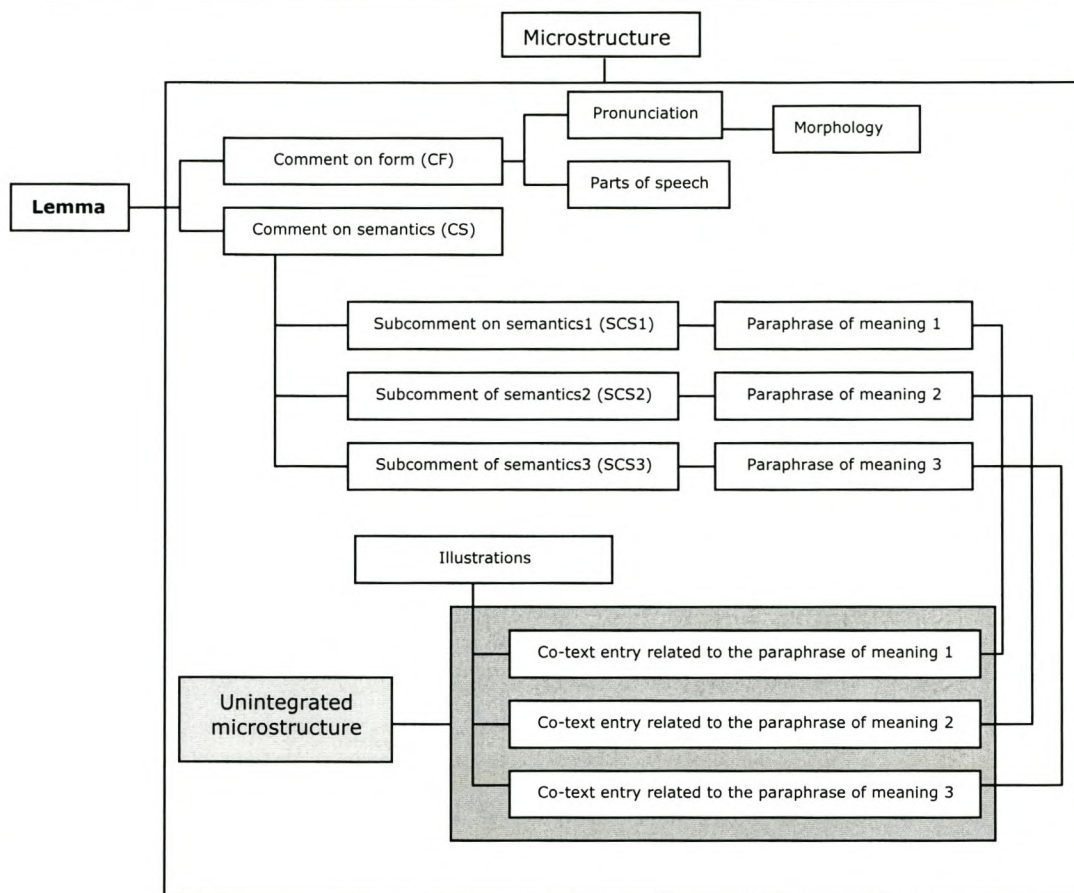
For the proposed dictionary model it will be difficult to make a clear distinction between comment on form and comment on semantics because of the way different categories of data will be included in the dictionary articles. In fact, the major lexicographic treatment will be made not on lemmata, as it is usually done in most dictionaries, but on equivalent items. In other words, the proposed dictionary model will have a mixed version of the microstructure, i.e. comment on form and comment on semantics. The lexicographical treatment will concern both, lemmata and equivalent items, comment on form and comment

on semantics will be displayed for each of them, as it will be demonstrated in further illustrations.

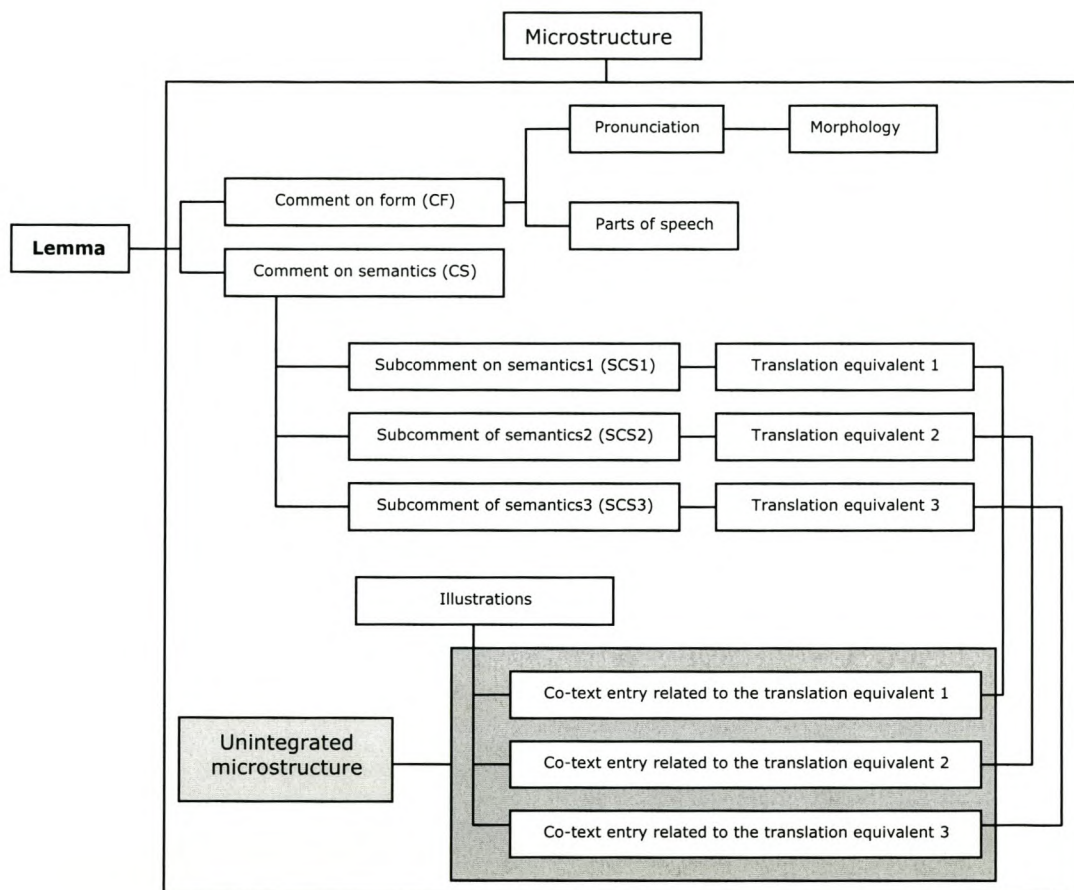
4.2.2.3.3.3. Unintegrated microstructure

For Hausmann and Wiegand (1989: 155) an unintegrated microstructure means that all semantic items, which are addressed at the lemma, appear in the first semantic subcomment of the article, so that this is a comment on the lexical meaning. And all example items and items are distributed in the semantic subcomment coming after the first one. Gouws (2001b: 87) specifies that an unintegrated microstructure displays addressing between a co-text entry and the relevant paraphrase of meaning/translation equivalent. In the case of a lemma sign representing a polysemous lexical, a bilingual dictionary will give all translation equivalents and then present the co-text entries. Gouws remarks that for a polysemous lemma in a bilingual dictionary, the user could experience problems in attempts to coordinate a specific co-text entry with a specific translation equivalent due to the indirect relation. Hausmann and Wiegand share this point of view and consider that even a well-informed user cannot use that approach as an inner rapid access structure in the case of polysemous items.

Schema of an unintegrated microstructure in a monolingual dictionary



Schema of an unintegrated microstructure in a bilingual dictionary

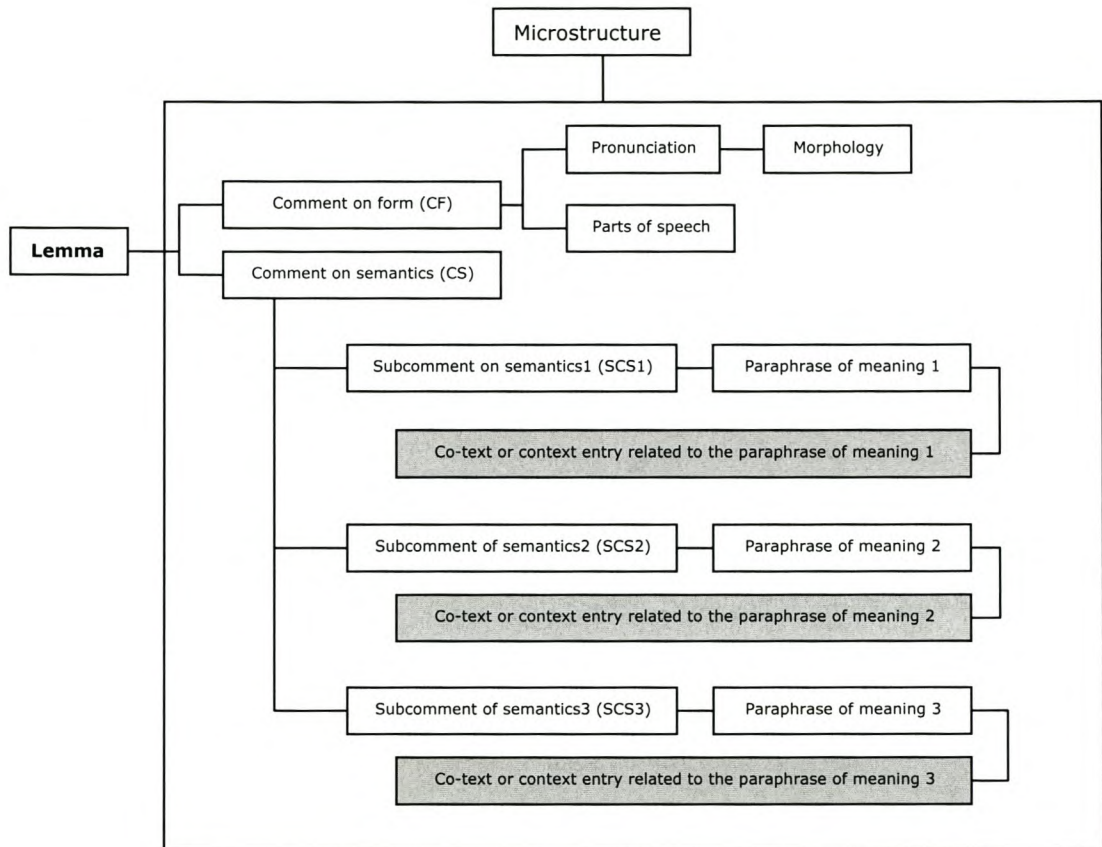


4.2.2.3.3.4. Integrated microstructure

Data concerning a lemma in an article can have different positions in a dictionary. Contrary to the unintegrated microstructure, an integrated microstructure displays a system of direct addressing between a paraphrase of meaning/translation equivalent and its co-text entry/entries. Each paraphrase of meaning/translation equivalent is immediately followed by the co-text entry illustrating the typical usage of the lexical item in question (Gouws, 1999b: 47). According to Hausmann and Wiegand (1989: 354), simple microstructures of general monolingual dictionaries may be called integrated when all items within the article which do not belong to the comment on form are located in the scope of a certain semantic item and belong to the same semantic subcomment to which the semantic item also belongs.

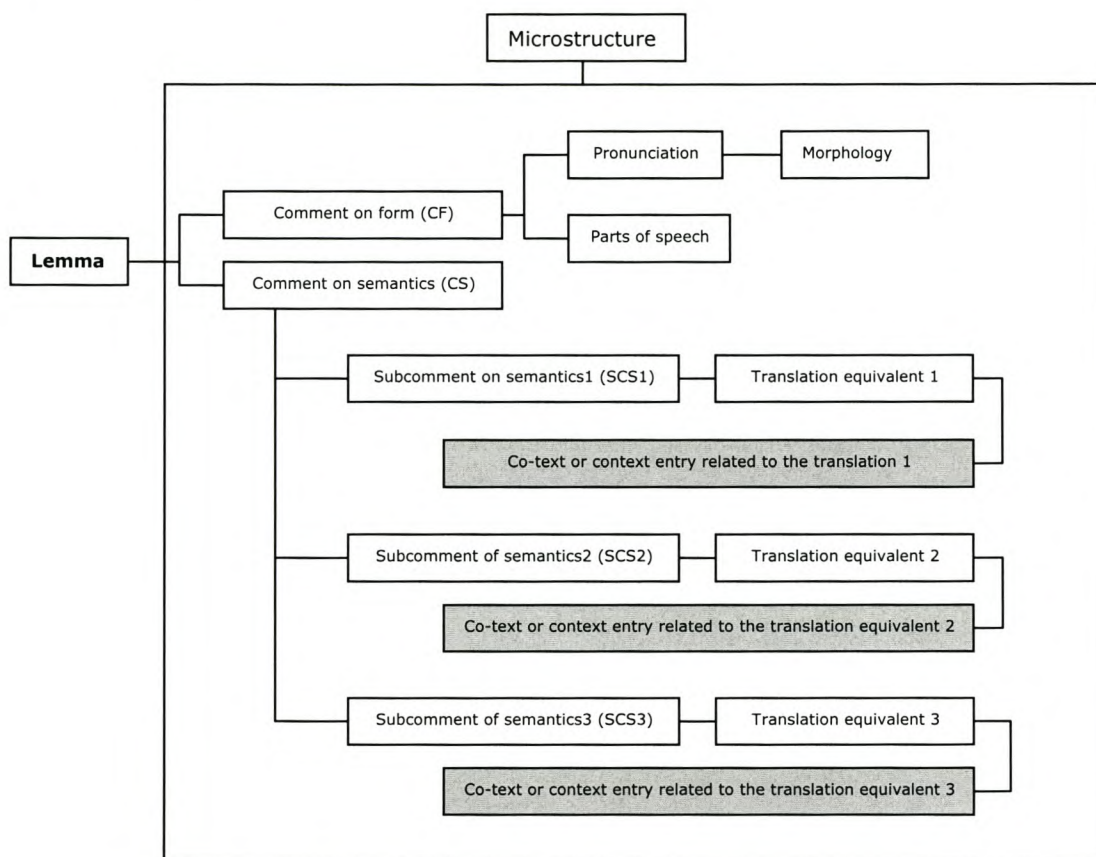
I do agree with Gouws (2001b: 87) who substantiates that the fact that no other occurrences of paraphrases of meaning/translation equivalents come between a given paraphrase of meaning/translation equivalent and its co-text entry decreases the textual condensation and makes it easier for a user to interpret the contents of the subcomment on semantics correctly. Especially in the treatment of a lexical item with many polysemous senses, the direct relation between co-text entry and paraphrase of meaning/translation equivalent ensures an optimal retrieval of information. This is a microstructure ideally suited for monolingual and bilingual pedagogical, desk/college and standard dictionaries.

Schema of an integrated microstructure in a monolingual dictionary



Each co-text or context entry is directly given after the related paraphrase of meaning. These co-text entries are considered as integrated microstructure.

Schema of an integrated microstructure in a bilingual dictionary

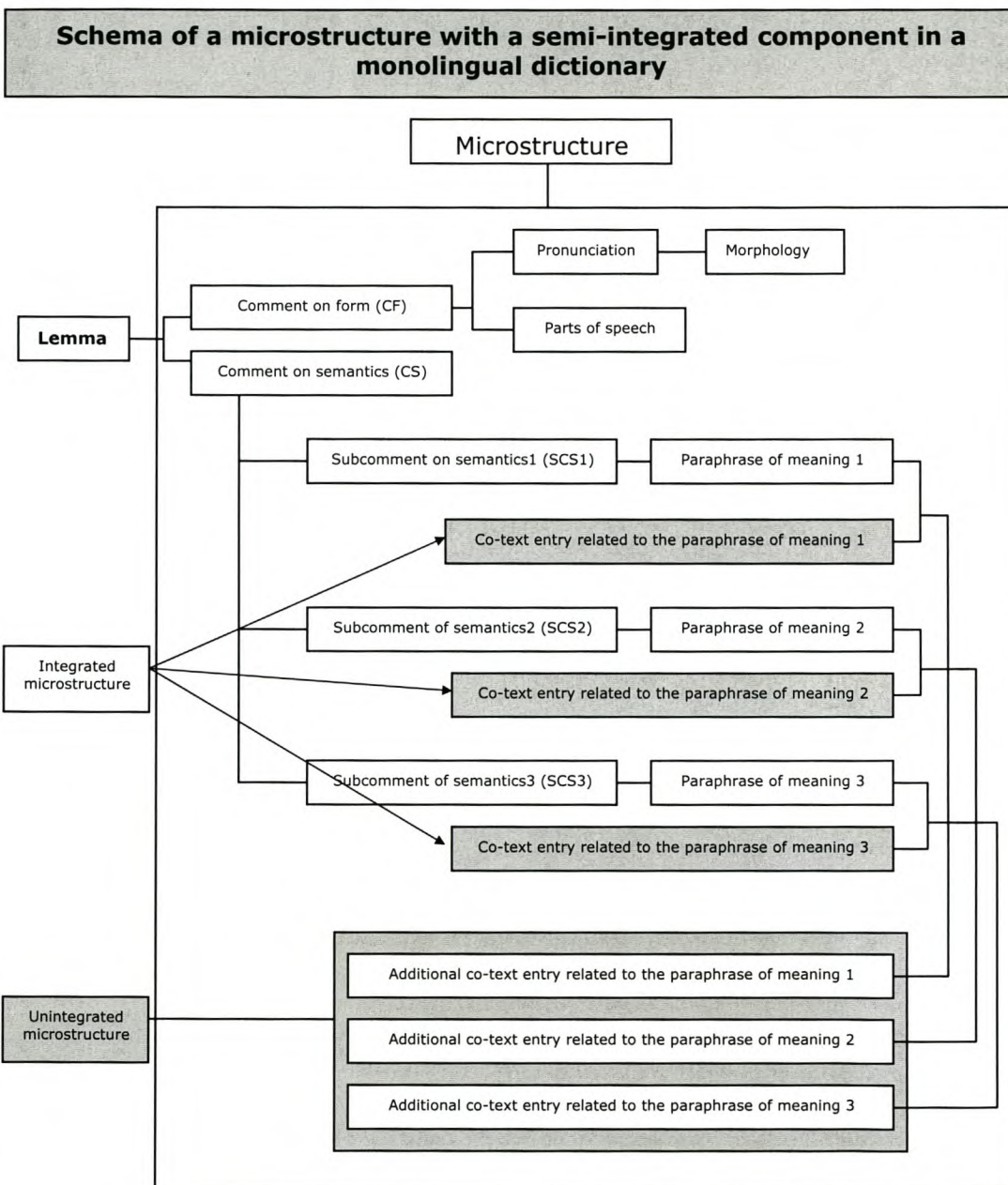


Each co-text or context entry is directly given after the related translation equivalent. These co-text or context entries are considered as integrated microstructure.

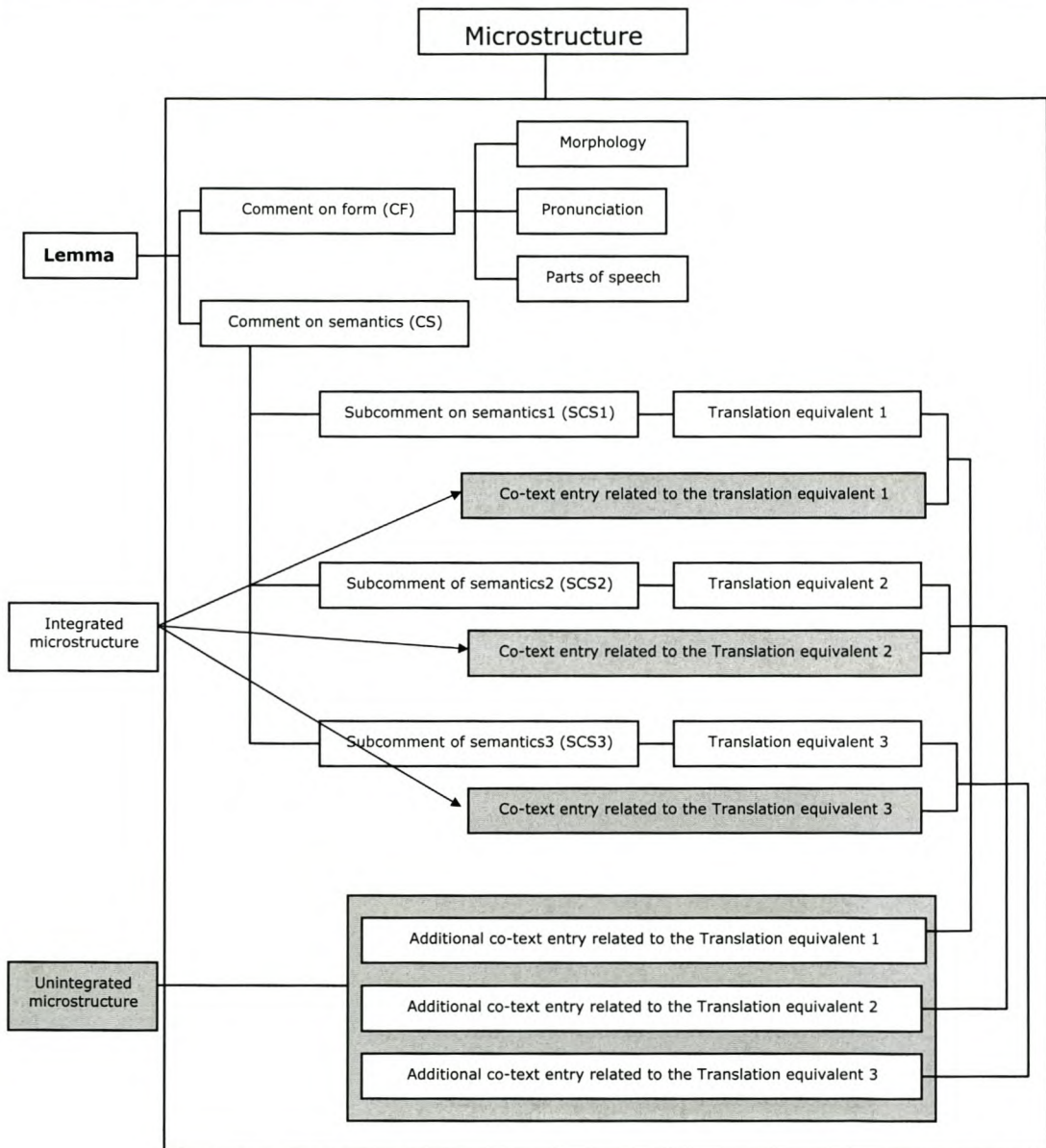
4.2.2.3.3.5. Semi-integrated microstructure

According to Gouws (2001b: 87) a semi-integrated microstructure is a hybrid form displaying features of both an integrated and an unintegrated microstructure and is typically used in more comprehensive dictionaries where lengthier articles with a variety of data types and search zones occur. He adds that in the semi-integrated microstructure the article structure has two distinct sections to deal with. In the first component (the integrated), a single co-text entry is added to each entry given as a paraphrase of meaning/translation equivalent. The second section (the unintegrated) is presented to accommodate additional co-text entries. In this section, the relation between each co-text entry and the relevant paraphrase of meaning/translation equivalent entry is explained by means of a clear and unambiguous cross-reference entry marking the co-text as addressed at a specific paraphrase of meaning/translation equivalent entry.

Gouws specifies that this type of microstructure assists the user with regard to both decoding and encoding needs. The integrated component is directed at a decoding function whereas the unintegrated component adds an encoding function to the article. A semi-integrated microstructure leads to a sophisticated lexicographic procedure and should be considered for more comprehensive bilingual and monolingual dictionaries.



Schema of a microstructure with a semi-integrated component in a bilingual dictionary



4.2.2.3.3.6. Rudimentary microstructure

In dictionaries, lemmata are not always treated at the same level. Indeed, in some dictionaries lexicographers also use a lexicographical procedure called **rudimentary microstructure**. According to Wiegand (1990: 56), a rudimentary microstructure prevails when the dictionary article is restricted to the comment on form. Thus the comment on semantics is substituted by a cross-reference entry. One also finds rudimentary microstructures in the case of lemma signs representing abbreviations. Where abbreviations are included as lemmata in the central list, their treatment is usually restricted to an entry indicating the full form of the abbreviation, (cf. Gouws, 1999b: 45).

The proposed dictionary model will not use *rudimentary microstructure*. As the intended model is a standard dictionary each presented lemma should have, at least, the two compulsory components of the microstructure.

4.2.2.3.4. Proposed microstructure

4.2.2.3.4.1. General remarks

According to Rey-Debove (1971:151) data directed at the lemma sign may be numerous, nevertheless not of equal usefulness. The microstructure has a constant organization, because the data categories given in the dictionary articles are more or less the same in each article. If the information program contains the types A, B, C, D (e.g. part of speech, definition, synonyms and homonyms) those four types will appear in each article. However, Rey-Debove specifies that certain data categories can be absent or omitted in some articles

according to the treated lemma sign. It is, for instance, evident that it would be difficult to give a synonym for each lemma, because it does not exist for every word in the language and even when it does, the lexicographer could judge it unnecessary to include such an entry in the article.

According to Hausmann and Wiegand (1989:346), lexicographers should identify in advance the **obligatory** or **compulsory microstructure**, i.e. the structure that is common to all articles. This obligatory microstructure contains those data categories, which have to be treated obligatory for each lemma sign. In a general monolingual dictionary the article of each lemma sign should contain a part of speech indicator and some form of semantic description – either a meaning paraphrase or a synonym. However, all articles do not include e.g. an entry representing an antonym of the lemma sign. An article slot for antonyms will only be utilized in articles where something more than an obligatory microstructure prevails. This constitutes an *extended obligatory microstructure* because it includes more data categories than the prescribed minimum (Gouws, 1999:45).

The microstructure I am proposing for my model will include all types of data categories that can be useful for the user. It will mainly be an integrated microstructure, with a semi-integrated dimension. This means that co-text entries will directly follow the paraphrase of meaning/translation equivalents. Some additional co-text data could be added at the end of the article according to the treated lemma. However, the lexicographer must make sure that the inclusion of this

co-text data will not make the articles unnecessarily complex (Gouws, 2002).

Practically, the microstructure for the model I propose will be divided into three major components, two obligatory and one optional:

- The first obligatory component will essentially display the comment on form and some aspects of the comment on semantics of the lemma. The comment on form will display what Hausmann & Wiegand call synchronic identification of the lemma (i.e. the phonetic pronunciation & the part of speech), and the comment on semantics will display the explanatory data (i.e. paraphrase of meaning of the lemma in French). It is important to specify that the comment on semantics in French should not be made for every lemma, but only for those ones that are rarely used in everyday conversations of the target users, and also for the lemmata that can have a specific cultural meaning in the Gabonese context. This approach is motivated by the fact that many young people do not always know the real meaning of many French words they use and as every language in the world; French has a local variety⁷ in Gabon. All these specificities have to be taken into account by the lexicographer.
- The second obligatory component of the dictionary model (the most important), entirely in Fang, will display in bold the equivalent item of the lemma, followed by its comment on form and comment on semantics. In fact, each Fang equivalent item

⁷We call that local variety the French of Gabon.

will be treated in the way lemmata are treated in monolingual dictionaries, i.e. with data displaying its synchronic identification (e.g. the phonetic pronunciation, the part of speech & the plural form) and its explanatory data mainly represented by the paraphrase of meaning of the equivalent item. That component will also display illustrations and examples related to the equivalent item, and will be complemented by some optional entries representing antonyms, synonyms or homonyms related to the Fang equivalent item. A cross-reference of the optional entry could also be given when that item is treated elsewhere in the dictionary.

▪ The third and last component in French will depend on the treated Fang equivalent item. That component will constitute the **extended obligatory** or **compulsory microstructure** because it will include more data categories than the prescribed minimum. It will display additional explanatory features or cultural data related to the Fang equivalent item such as encyclopaedic or ethnological data. This component will be written in French because the intended target users are not yet able to read fluently in Fang. It will be easier for them to retrieve the cultural information related to the equivalent item if it is written in French.

The dictionary model I propose has the purpose to help with the standardization of Fang. The innovative approach of the proposed model will be that the main lexicographical treatment will be made in

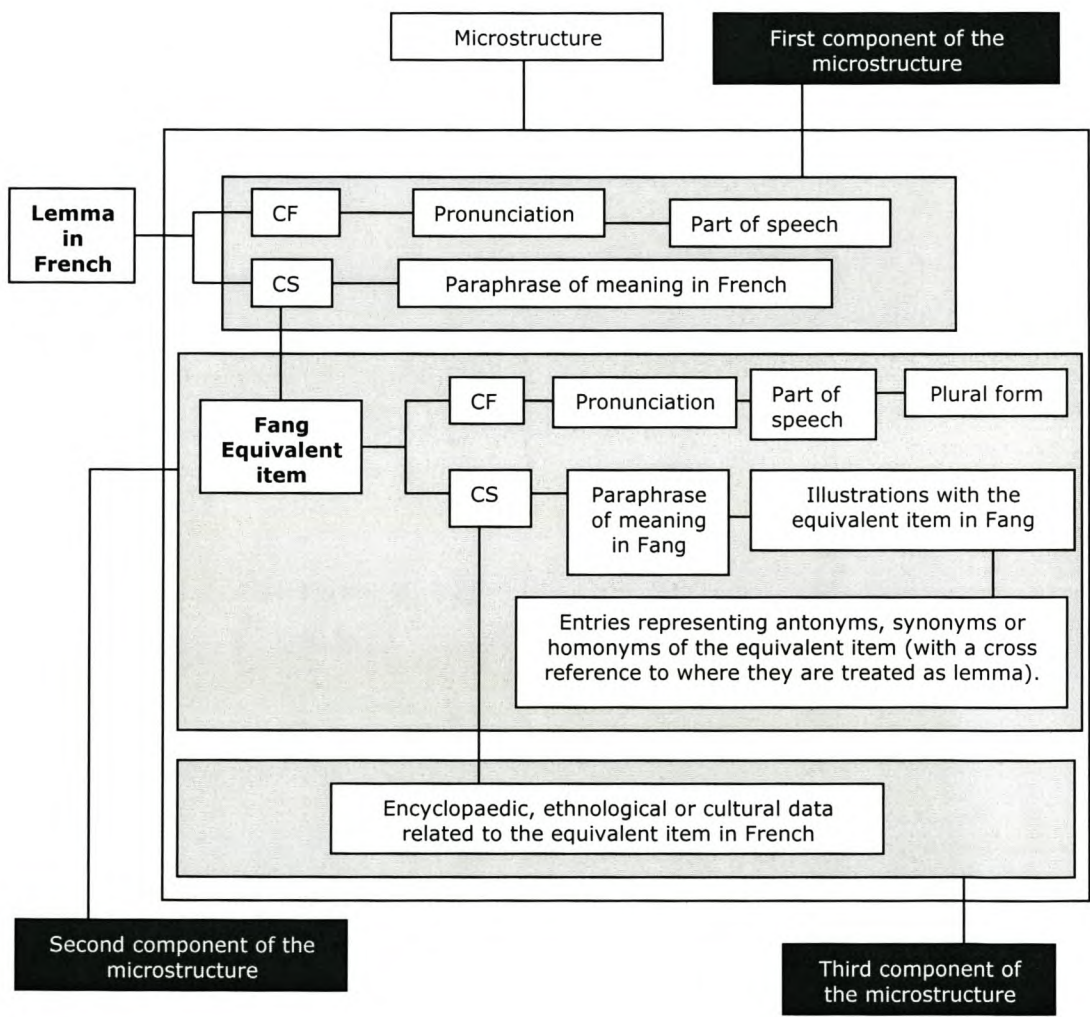
the target language. In that regard, it can be noted that, an entire lexicographical treatment of the equivalent item is made in Fang. The main aim of that approach is to accustom the users with the language and the writing system. Somehow that dictionary model will have a monolingual dimension that could facilitate the compilation of monolingual dictionaries in Fang. The presentation of lemmata and their treatment in the first component, as well as the presentation of cultural data in the third component in French is only motivated by the fact that French can serve as easier search route for the target users. Since the target users are accustomed with French, it will be easier for them to have access to the needed information via French.

4.2.2.3.4.2. Micro-architecture

Apart from the access structure, the ways data categories are organized or structured also play the role of search zone in dictionary articles as that organization allows the user to easily access the needed information. For instance, if three data categories (paraphrase of meaning, translation equivalent, and cultural data) have to be included in the treatment of a lemma, the lexicographer has to make provision for three article slots or search zones in the dictionary article. This partly corresponds to Wiegand's (1996d) concept of **micro-architecture**. Some non-typographical markers (e.g. arrows, italics, bold, etc.) introducing different data categories can also represent the micro-architecture of a dictionary article. The micro-architecture allows a clear presentation of data categories in a dictionary article in the way

that these data categories can easily be distinguished or identified by the user.

The described microstructure can be schematized as below



The following illustrations are practical examples of how articles will look like in the proposed dictionary model.

danse [dās] n.f. Suite de mouvement rythmique du corps, évolution à pas réglés, le plus souvent à la cadence de la musique ou de la voix.

abok [abok] e.d. Ab. **mebok**. Ngura aval ya nfog nyol eyong azik ngëki dzya dalqt. *Ku e këlë anë abok binëga.*

dot [dot] n.f. Biens donnés par un tiers dans le contrat de mariage.

nsua [nswa] e.d. Ab. **misua**. Mibeng (monë ya byëm) fam da kë bëkya beny eyong avëk minëga. *Nsua Ndong a nga mvëk ngal obë mintët mëwom mëtan.* ⇔ **bili bi kama**²⁴¹.

+ Dans la tradition Africaine en general, et Fang en particulier, la dot est un ensemble de biens matériels et financiers qu'un homme apporte à la famille de sa fiancée pour sceller traditionnellement leur mariage. La cérémonie a souvent lieu dans le village de la femme en présence des deux familles. Elle se déroule en terme de négociations, chaque camp usant de stratégies et de verbes: le famille de l'homme voulant donner le moins possible, et celle de la femme voulant gagner le plus possible. Les deux fiancés n'interviennent généralement pas (mais ils peuvent être consultés en privé), toute la procédure est laissée aux parents, jusqu'au moment où les parents de la fiancée demandent à cette dernière de prendre l'argent proposé par les parents de l'homme. Si cette dernière rend l'argent aux parents de son prétendant, cela veut dire qu'elle ne veut pas aller en mariage ou du moins épouser l'homme qui demande sa main. Mais si elle donne l'argent à ses parents, cela veut dire qu'elle accepte le mariage.

NB: Abbreviations & symbols in the articles

n.f.: Indication of part of speech in French

e.d.: Indication of part of speech in Fang

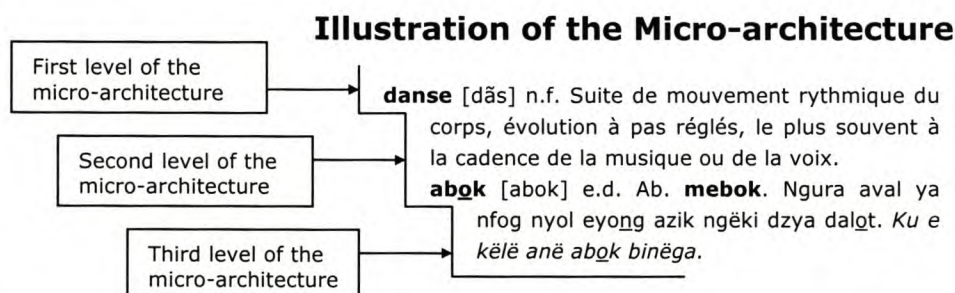
Ab.: Introduce the plural form in Fang

⇔ : Introduce a cross reference

+: Introduce cultural data

In the proposed dictionary model the *micro-architecture* will correspond to the presentation of each component of the proposed microstructure. Indeed, each component of the microstructure will

start with a new paragraph in the way that it can be easily identified at the extreme left of the article. Another important microstructural aspect will concern the arrangement of each data category. The article will be presented as upside down stair case. The first level of the stair case will start with the lemma, the second level of the stair case will start just below the lemma up to the equivalent item, and the third level will start below the equivalent item. This presentation will allow the user to quickly identify each search zone in the article slot.



4.2.2.3.5. Detailed description of microstructural features

4.2.2.3.5.1. Introduction

As shown in the previous section, the proposed model will lead to a hybrid dictionary because it will include bilingual as well as monolingual features. The lexicographic treatment also includes the transfer of a number of types of semantic data. Besides translation equivalent(s), the type of translation dictionary I propose will contain a variety of data categories that will have an important place in the presentation of the microstructure. As the proposed model is a translation dictionary, an analysis of equivalent relations in bilingual dictionaries will be undertaken in the following lines. It should be

noted that in that analysis I will only be focusing on the denotation aspect, and not on register and style.

4.2.2.3.5.2. Equivalent relations in bilingual dictionaries

4.2.2.3.5.2.1. General remarks

As regard the presentation of translation equivalents, the treatment of a lexical unit compels the lexicographer to give a target language entry with the same semantic value. However, the user should also know whether the translation equivalent can be used in the same register and pragmatic context as the source language item (Gouws, 1996:72). The task of the bilingual dictionary is to provide words and expressions in the source language with counterparts in the target language which are as near as possible, semantically and as regard style level i.e. register (Svensén, 1993: 140).

A translation equivalent is a form in the target language, which can be used to substitute the source language form in certain contexts. The relation in an article of a translation dictionary between the lemma, that is the source language form, and the translation equivalent(s), the target language form(s), is known as an *equivalent relation* (Gouws, 2000: 3). The basic purpose of a bilingual dictionary is to coordinate with the lexical units of one language those lexical units of another language, which are equivalent in their meaning (Zgusta, 1971; 294). Usually one can find three main types of equivalent relations in bilingual dictionaries: **absolute**, **partial** and **surrogate** equivalence.

4.2.2.3.5.2.2. Absolute equivalence

When the meaning of a lexical unit in the source language is exactly the same as the meaning of another lexical unit in the target language, there is an **absolute** or **congruent** or **full** equivalence. Usually in absolute equivalence the source and the target items have the same meaning and they can be used in the same register. They do have the same usage restrictions. However, the fact that both source and target language items have the same meaning and usage restrictions does not only lead to a relation of full semantic equivalence but also full communicative equivalence. By giving translation equivalents the lexicographer empowers his/her users because he/she gets the necessary information to communicate in the target language within the same register as used in the source language (Gouws, 1996:72-73).

Demonstration

X= meaning

A= Source language

B= Target language

1= Lexical unit

←→ = If and only if (when)

Equation: $1A = 1B \leftarrow\rightarrow XA = XB$

Explanation: The lexical unit in language A (i.e. source language) is equivalent to the lexical unit in language B (i.e. target language) if and only if the meaning of the lexical unit in language A is equivalent to the meaning of the lexical unit in language B.

Example:

The French term "sagaie" (i.e. assegai), described by the *Dictionnaire Universel* as "Javelot dont une extrémité est munie d'un fer de lance ou d'une arête de poisson" (i.e. javelin with an extremity equipped with a pointed iron or fish bone) will have a full equivalence relation with the Fang item "akong", because both items have the same and entire meaning.

4.2.2.3.5.2.3. Partial equivalence

It is generally accepted that full equivalence has a less frequent occurrence in bilingual dictionaries than partial equivalence. This also applies to the relation of equivalence holding between source language lexical items and their target language equivalents. The fact that most lexical items are often marked for usage in certain specific registers and the consequent stylistic restrictions can be regarded as some of the major reasons why the co-ordination of items between two languages often only succeeds partially (Gouws, 1996:75). When the meaning of source language items refers to one part or aspect of the meaning of the target language one (or vice versa), one can speak of **partial equivalence**. In the partial equivalence relation, the source language item can have more than one equivalent in the target language, however its equivalence relation with each target language item is only partial and not full. In that case one can speak of **divergent equivalence**.

Demonstrations**Partial equivalence**

To an item X in the source language A corresponds an aspect of the item Y in the language B.

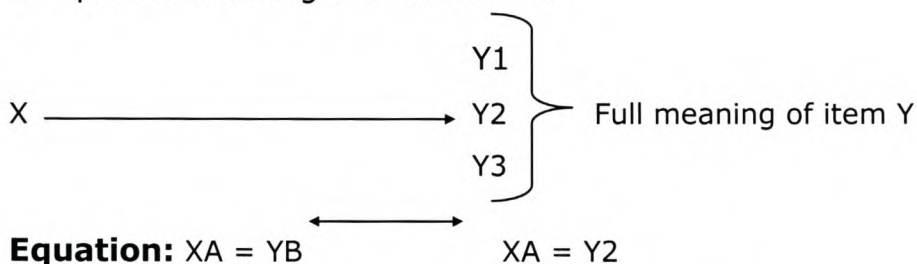
X = item in source language A

Y = item in the target language B

1 = partial meaning 1 of the item Y

2 = partial meaning 2 of the item Y

3 = partial meaning 3 of the item Y



Explanation: The item X in language A (i.e. source language) is equivalent to the item Y in language B (i.e. target language) if and only if the meaning of the item X in language A is equivalent to the partial meaning 2 of the item Y in language B.

Example:

The French term "père" (i.e. father), described by the *Dictionnaire Universel* as homme qui a engendré un ou plusieurs enfants (i.e. man who has fathered one or many children), has a partial equivalence with the Fang item "esa", which has two senses in Fang: father and brother of the father.

Divergent equivalence

To an item X in the source language A corresponds an aspect of the items Y, Z and M in the language B.

X = item in source language

Y = item 1 in the target language

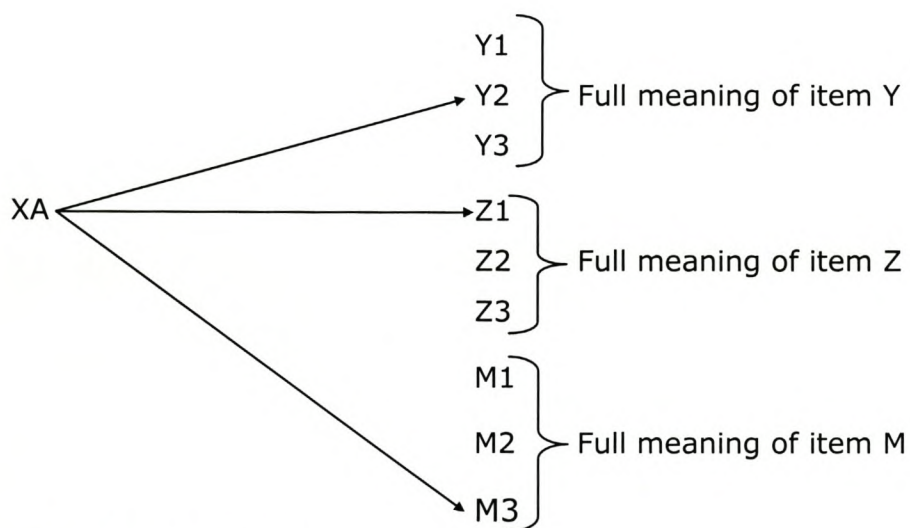
Z = item 2 in the target language

M = item 3 in the target language

1 = partial meaning 1 of the items Y, Z and M.

2 = partial meaning 2 of the items Y, Z and M.

3 = partial meaning 3 of the items Y, Z and M.



Y2= sense 1 of XA

Z1= sense 2 of XA

M3=sense 3 of XA

Explanation: The lexical unit X in language A (i.e. source language) is equivalent to the lexical units Y, Z & M in language B (i.e. target language) if and only if the meaning of the lexical unit X in language A is equivalent to the partial meaning Y₂, Z₁ & M₃ of the lexical units Y, Z & M in language B.

Example

The French term "oncle" (i.e. uncle) defined by the Larousse dictionary as "le frère du père ou de la mère" (i.e. brother of the father or mother) will have two equivalent items in Fang; "esa" for brother of the father, and "nyandome" for brother of the mother.

4.2.2.3.5.2.4. Surrogate equivalence

Besides absolute and partial equivalences in bilingual dictionaries, there is a third type of relation between the source and the target language. This relation is established in an article where the target language has no lexical item to be used as translation equivalent of the lemma (Gouws, 2001:21). The lack of suitable translation equivalents in the target language can indicate a cultural gap between the language pair. The absence of target language equivalents compels the lexicographer to employ alternative strategies in his quest to ensure semantic and communicative equivalence between source and target language (Gouws, 1996:81). In that situation one can speak of **surrogate** or **substitutive** equivalence.

For the dictionary model I propose, when confronted with the lack of equivalence between French and Fang, the lexicographer can adopt three types of surrogate strategies or approaches to ensure

semantic and communicative equivalence between the source and the target language:

1. In some cases, the lexicographer can borrow words from other languages when they are already known and used by Fang people with the same semantic value and the same register of communication. That strategy has been used in many languages, even in Fang, to name new realities unknown before. For instance Fang people have borrowed from English terms to name realities like *soap* (*sob* in Fang), *towel* (*tawël* in Fang), or *motorcar* (*mëtwä* in Fang).

2. Another approach could be to create substitute terms either according to the physical description of the object to be named, by comparing the reality to be named with other realities, or according to the function of the reality or object to be named. This can be seen in existing Fang words. For instance the object *train* is designed in Fang by the group of words *misini a mikong*, which literally means *bicycle of frogs*, because physically the train coaches look like a queue of frogs, and the train wheels are like bicycle wheels when the train is in motion. In the same vein, the *stereo radio* is called *ewala mikob* (i.e. box of words), the church *nda nzam* (i.e. God house), etc.

3. The third and last strategy, the most strongly recommended one, will be for the lexicographer to initiate an investigation through the language speakers and ask them how they can intuitively name the new reality. If the lexicographer uses a live recording investigation,

he should have with him photos or physical representations of realities to be named. In case of an investigation with forms to be completed by speakers, photos or images of things to be named should go with these forms. That type of investigation will allow the lexicographer to have a broad idea of how the speakers would like to name the realities or things in question. That descriptive approach, contrary to the prescriptive⁸ one, allows the lexicographer to choose terms that will be easily accepted and used by the speakers.

4.2.2.3.5.2.5. Remarks on translation

The major task of a bilingual lexicographer is to find appropriate equivalents in the target language to the units of the source language. This task involves a great deal of translation (Al-Kasimi, 1977:59). Jakobson (as cited by Al-Kasimi, 1977:59) labels three different kinds of translation:

1. Intralingual translation or rewording is an interpretation of verbal signs by means of other signs of the same language.
2. Interlingual translation or translation proper is an interpretation of verbal signs by means of some other language.
3. Intersemiotic translation or transmutation is an interpretation of verbal signs by means of signs of nonverbal sign systems.

According to Al-Kasimi (1977:59), "bilingual lexicography is concerned with the translation proper or interlingual translation". He adds that

⁸ Usually with a prescriptive approach the lexicographer chooses terms that he/she thinks better for the language. However, there is a risk that the speakers can reject the terms chosen by the lexicographer, because they are not familiar with them.

"one of the linguistic processes to establish translation equivalents is the commutation test, which makes use of a bilingual as an informant. The bilingual is presented with a sentence in the source language and requested to put it in the target language. Then the items of the sentence are gradually changed one by one and every time an item is changed the bilingual provides the proper translation". Nida (as cited by Al-Kasimi, 1977:60) recommends that even if a lexicographer has a long experience with the foreign language, he/she should seek the constant help of a native informant of that language in order to weed out improper usage and avoid translationisms. Al-Kasimi makes another important remark as far as translation in a bilingual dictionary is concerned. According to him, the translations of entry words are usually of two types: translational equivalents, and (b) explanatory equivalents.

(a) A translation equivalent is a lexical unit which can immediately be inserted into a sentence in the target language (e.g. in an English-French dictionary boy = garçon).

(b) An explanatory equivalent or descriptive equivalent is one that cannot always be inserted into a sentence in the target language. To give a translation equivalent for the words "boyhood", the lexicographer may give "adolescent" or "jeunesse". But the English entry word is restricted to male children whereas the two French translational equivalents are not. The lexicographer may try a compromise and say for instance "boyhood" = adolescence (d'un garçon).

All these examples show how difficult it is sometimes for the lexicographer to find the right equivalent in the target language. That situation is caused in particular by the cultural gap between the members of the language pair and the scientific and technological terminology, which does not exist yet in local languages of developing countries. According to Al-Kasimi (1977:61) the UNESCO committee that studied the use of vernacular languages in education pointed out five different ways to expand or extend a vocabulary: (1) word borrowing, (2) coinage, (3) giving new meaning to existing words, (4) extending the meaning of existing words, and (5) compounding new words from existing elements from the language and some other ones. The committee also recommended that devices (3), (4) and (5) should be preferred to (1) and (2). It also said that borrowed words should be adapted to the sound system and grammar of the language, and that borrowed scientific terms should be consistent as to the type of information and language origin. The UNESCO committee also advised that the acceptability of all new words should first be tested before they are finally adopted. In my view, lexicographers when compiling bilingual dictionaries, particularly in developing languages such as Fang, should take all these recommendations into account.

4.2.2.3.5.3. Addressing structure

4.2.2.3.5.3.1. General remarks

According to the new concept of the dictionary structure proposed by Hausmann and Wiegand (1989: 349), the microstructure of the dictionary articles is not the only (partial) structure within the

complete article structure. The other important structure is the addressing structure. Hausmann and Wiegand (1989: 129) specify that the addressing structure is the structure, which describes all types of addressing in a given dictionary. Gouws (2001b: 89) adds that each microstructural entry is part of the treatment of the lemma sign of the given article or a form of treatment or lexicographic comment on another microstructural entry elsewhere in the dictionary. Microstructural entries are directed or addressed at specific targets. The addressing structure of a dictionary is the system according to which these procedures of one entry directed at another are employed. Hausmann and Wiegand (1989: 349) complete this point of view by saying that each item refers to an addressee through an address given in the article. Normally the central address of a dictionary article is the item giving the form of the lemma sign and consequently the lemma. When the address is the lemma, we have *lemmatic addressing*; when it is the sublemma, we have *sublemmatic addressing*; in all other cases the items are addressed by *non-lemmatic addressing*. If all items inside the article are addressed at the lemma, we have *full lemmatic addressing*.

4.2.2.3.5.3.2. Lemmatic addressing

According to Gouws (2001b: 89) the lemma is the main treatment unit in an article and therefore constitutes the most typical address in that article. He adds that the lemmatic addressing is a procedure where the lemma is the address of an entry. In a dictionary, which adheres to a strict initial alphabetical ordering, all the lemmata would be arranged

vertically and each lemma will be the guiding element of an article. Lemmatic addressing will always have one of these lemmata as an address.

For the dictionary model I propose, the microstructure will display a lemmatic addressing because some data such as the pronunciation, the part of speech and the paraphrase of meaning in French will be addressed at the lemma. However, that lemmatic procedure will only serve as guiding element that will direct the user through the main address, i.e. the equivalent item.

4.2.2.3.5.3.3. Sublemmatic addressing

In a dictionary where a sinuous lemma file prevails, the occurrence of niched and nested lemma clusters will lead to a situation where all lemmata are presented as sublemmata attached to article stretches. Members of a niched or nested cluster may also be the address of an entry. This addressing procedure is known as sublemmatic addressing. The sublemmata remain part of the macrostructure of the dictionary; therefore sublemmatic addressing is a type of lemmatic addressing.

For the model I propose, since the dictionary will not include sublemmata, there will be no data addressed at sublemmata, and consequently the dictionary will not display any sublemmatic addressing.

4.2.2.3.5.3.4. Full lemmatic addressing

When all items inside the article are addressed at the lemma, we have full lemmatic addressing. This means that there is no topic switching because all lexicographic statements comment on the lemma sign

(Hausmann and Wiegand, 1989: 349). In the situation of full lemmatic addressing in bilingual dictionaries, equivalent items appear without any treatment addressed at them. However, sublemmatic addressing can be part of the full lemmatic addressing if it is regarded as a type of lemmatic addressing.

For the model I propose, the main lexicographical treatment will be addressed at the equivalent item, this means that all the items inside the article are not exclusively addressed at the lemma. Consequently, the microstructure will not display a full lemmatic addressing.

4.2.2.3.5.3.5. Non-lemmatic addressing

For Hausmann and Wiegand (1989: 349) non-lemmatic addresses are addresses, which belong exclusively to the microstructure. They may be called subaddresses. Gouws (2001: 90) adds that non-lemmatic addressing is a procedure involving an entry not functioning as lemma, as an address. The address is the topic of the specific treatment procedure. The use of non-lemmatic addressing implies a system of topic switching because each non-lemmatic address is a new topic within the article.

The dictionary model I propose will essentially utilize a non-lemmatic addressing procedure. The main lexicographic treatment will be addressed at the equivalent item and not at the lemma. Indeed, for a successful retrieval of the information needed by the user, a system of topic switching is necessary. Addressing procedures will not be directed at the macrostructural domain, but they will mainly involve

the microstructure of the dictionary. This is pertinent because each non-lemmatic address becomes a treatment unit or new topic within the article (cf. Hausmann and Wiegand 1989:329). A system of topic switching in the proposed model will essentially concern translation equivalents and all data categories addressed at these translation equivalents.

4.2.2.3.5.3.6. Remarks on addressing structure

The dictionary model I propose will only contain two types of addressing:

1. Lemmatic addressing, because an important place will be given to the treatment of the lemma sign (pronunciation, part of speech, paraphrase of meaning and item equivalent).
2. Non-lemmatic addressing, because each equivalent item will also receive a full treatment (pronunciation, meaning paraphrase and examples). In some cases, encyclopaedic, cultural or ethnological entries addressed at the equivalent item could also be given.

Sublemmatic and full lemmatic addressing will not be part of the model I propose. The dictionary will present a strict alphabetic macrostructure, which means all potential sublemmata will be presented as lemmata. To be clearer, the dictionary model I propose will not have sublemmata, and consequently no sublemmatic addressing will take place. As far as full lemmatic addressing is concerned, because the dictionary model I propose is a standard bilingual dictionary, co-text entries will not only be addressed at the

lemma (in French), but also at the equivalent item (in Fang). That means it will not be possible to find in the dictionary any article in which the lemma is the only address. This type of addressing procedure is also called direct addressing integrated microstructure.

4.2.2.3.6. Microstructure and cultural data

Different dictionaries can have different functions. The main distinction with regard to dictionary functions is that between communication-directed functions and knowledge-directed functions. Communication-directed functions should help the user in solving problems in the sphere of communication whereas knowledge-directed functions aim to increase the user's knowledge regarding a specific topic. The selection, the nature and the presentation of data for any given dictionary may never be done in a random way but should be determined by, among other things, the function(s) of the dictionary, the dictionary type and the needs and reference skills of the target user of the dictionary (Gouws, 2002:56).

As the proposed dictionary model will be used in a multilingual environment, it will function not only as a linguistic instrument but also as a cultural instrument. Therefore, the data type to be included in the French-Fang standard translation dictionary should not focus exclusively on linguistic data, but also on cultural data. Functionally, the inclusion of cultural data will be relevant for the proposed model in the way that it will help young Fang people to improve their cultural knowledge about Fang. In fact, the majority of target users of the proposed dictionary are not too familiar with the culture of people. The

way cultural data should be presented and treated should enhance the quality of their knowledge about Fang people and their cultural particularities. Gouws (2002:63) encourages such an approach by saying that "lexicographers compiling dictionaries to be used within a multilingual and multicultural environment should pay serious attention to the inclusion of cultural data in their dictionaries". He adds that "in all types of general bilingual dictionaries the inclusion of cultural data should be regarded as compulsory and lexicographers should negotiate various options to present it in the best manner".

"Many dictionaries have employed the back matter section to include a variety of texts" (Gouws, 2002: 57). For the proposed model I have opted to include cultural data in the treatment presented in the central list. This means that cultural data will be accommodated within the comment on semantics of the equivalent item. In the articles where cultural data will be added, an additional search zone will be introduced to complement the structure of the default article (i.e. the obligatory components of the microstructure). A specific structural marker (+) will mark the search zone accommodating the cultural data in order to allow an easier retrieval of information by the users. Mainly, the treatment presented in the cultural data search zone will include encyclopaedic and ethnologic explanations related to the Fang equivalent item. These types of articles, which belong to the category of complex articles (cf. Gouws, 2002: 65) in combination with outer texts, can be regarded as an ideal way to present and treat cultural data in bilingual dictionaries. Therefore, in the proposed model the data exposure of the dictionary will employ both article stretch internal

and outer text procedures to make the target users aware of this data category. I entirely agree with Gouws (1999 as cited by Gouws, 2002:66) who thinks that lexicographers working with complex articles will have to employ an extended obligatory microstructure, i.e. a structure that allows the additional data types allocated to a specific article type. That extended microstructure corresponds to the third component of the microstructure in my proposed model.

4.2.3. Data distribution structure in outer texts

4.2.3.1. Introduction

Successful dictionary use relies on a dictionary text, which allows easy access to the needed data. Besides the central list, a dictionary should have other components, which have to contain an explanation on the editorial system as well as data added to complement the central list (Gouws, 2000: 21). This additional data is presented in the texts outside the central list (that is why it is called outer texts), either before the central list (*front matter texts*) or after the central list (*back matter texts*). The use of outer texts allows the lexicographer to include entries that will not typically appear in a linguistic dictionary. It is important to make a clear distinction between *integrated* and *unintegrated outer texts*. Unintegrated outer texts complement the central list and are not needed to retrieve the information presented in the articles of the central list. Integrated outer texts function in coordination with the central list and are needed to ensure an optimal and full retrieval of information (Gouws 2001b: 89).

4.2.3.2. Front matter texts

4.2.3.2.1. General remarks

In most dictionaries, the front matter is the section in which the explanation regarding the editorial system is presented. That will also be the case for the model I propose. Indeed all the dictionary components (different structures) will be explained, particularly the article components as well as all symbols, icons, characters or signs used. That section will also give guiding notes on pronunciation in French and Fang, as well as abbreviations. Even if the majority of entries in the front matter will be helping the users to understand abbreviations and symbols used in the central list, they will primarily be functioning as unintegrated outer texts because they will not directly be regarded as complementary data of the treatment made in the central list. They will present other types of data in their own right in order to help the user to an optimal and full retrieval of information.

In active bilingual dictionaries, the front matter often also includes lists of principal parts of irregular verbs. For the proposed model, the front matter will also include the most used verbs (regular and irregular) in Fang with their conjugation. The presentation of these conjugated verbs will play a communication-directed function because the aim is to help the user for the production of texts in Fang. These types of entries in the front matter texts are motivated by target user situations. In fact, young Fang people need to communicate in the language, not only between themselves, but also with old Fang people who do not speak French. The front matter of the proposed dictionary will intervene indirectly when they will have some communicational

problem in terms of constructing sentences. Verbs are essential elements in sentences, knowing them or how they have to be conjugated in different times, can solve many problems of communication for users. This could be achieved by consulting the front matter texts of the intended dictionary.

4.2.3.2.2. Front matter and user's guidelines

When prospective users are contemplating buying a certain dictionary, they need a coherent account of its general features. Throughout the front matter of the dictionary, they can find a section that should help in the proper use of the book. The section where this type of data is given in the front matter texts is called the users' guidelines of the dictionary. The concerned data should provide information about the general purpose of the dictionary, the intended user group, the general organization of the dictionary, the scope of the dictionary, etc. All these aspects should be broadly explained in a way understandable to laypersons. A dictionary user's guidelines should also provide a kind of "description of the goods", addressed mainly at experts, in particular to those ones who have to decide whether the book is suitable for their needs (e.g. teachers in schools). This description should give a more technical account of the design and approach, the principles of selection, and the sources used. However, the main aim of the users' guidelines is give to the user instructions for a well and efficient utilization of the dictionary. The instructions have to be made in a way that will be clear and accessible to the user. These instructions should include an explanatory description of the macro-

and microstructure, and the data categories included in the dictionary (Svensén, 1993:230).

4.2.3.3. Back matter texts

4.2.3.3.1. General remarks

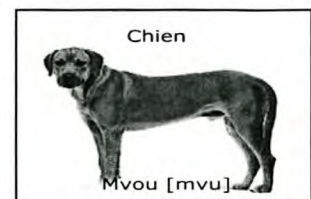
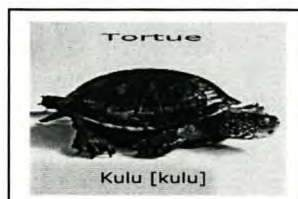
The back matter is generally the last part of the dictionary and it is the only section, which is not a compulsory component. Most of the time this section is used for the presentation of additional data that is not necessarily related to the data given in the central list. For the theoretical model I propose, the back matter will be divided into three main parts:

- The first one will present Fang proper and common names with some encyclopedic data (given in French) related to these names (e.g. Fang history). This section will allow the lexicographer to say more about Fang people, their history, their culture, their mythology and believes, etc. that are not always well known by the target group (e.g. the 'mvet'⁹). That section will play a knowledge-directed function in order to allow the user to know more about Fang people. That section could also be useful for academic purposes in future research of Fang.
- The second section will be presenting pictures (photos, icons, images, drawings, etc.) of certain domains or aspects that could not be given, in detail, in the central list (e.g. animals, fish, parts of body, trees, flowers, birds, etc.). Those pictures will be given

⁹ The 'mvet' is a collection of Fang epic tales that relates mysteries and adventures of Fang people. An initiated person who at the same time sings and plays a Fang musical instrument called 'mvet' only may tell those tales.

with their names in French (on the top of the picture), as well as their equivalents and pronunciation in Fang (on the bottom of the picture). These pictures will play both a knowledge-directed function (pictures and names in Fang) and a communication-directed function (pronunciation). They will not necessarily represent things named in the everyday conversation of the target group, but could be important for their general background or for their academic purposes. The main goal of presenting those pictures is to give the user the opportunity to make an easy association between images and names in Fang. To a specific picture will correspond a specific name in Fang and that will be easier for them to transfer their knowledge from the book (i.e. the dictionary) to the reality. It is important to specify that as lemmata in the central list, names of representations in pictures will be written in the standard alphabet and the guidance on their pronunciation should be done according to the way they are pronounced in the standard dialect. The purpose of the proposed dictionary model, once again, is to help with the standardization of Fang.

Examples



- The last section of the back matter, will present some additional data, such as Fang dialects and their geographic location in Gabon and in Africa, etc. These data will mainly have a knowledge-directed function, because they have the purpose to give the user more general information about the language.

The back matter will present both unintegrated and integrated outer texts. Unintegrated because some data will not be linked directly to the central list, and integrated because some cultural data will be regarded as complementary to the treatment given to some lemmata or equivalent items in the central list.

4.2.3.3.2. Back matter and complex articles

“The introduction of the notion of synopsis articles created the opportunity for lexicographers to negotiate a heterogeneous article structure when planning a new dictionary project. In a synopsis article the lexicographer treats the relevant lemma, but also gives data directed at other lemma signs included elsewhere in the dictionary. It could result in a kind of overview of the field to which the specific lemma belongs. Synopsis articles often tend to be more encyclopaedic than their single article counterparts” (Gouws, 2003:64). In monolingual and bilingual dictionaries a distinction can be made between single articles, which will represent the default article type, and complex articles in which a more comprehensive treatment is given. A complex article is not necessarily a synopsis article, but a synopsis article belongs to the category of complex article (cf. Gouws, 2003: 64-65).

Despite the fact that it is recommended to include a limited number of complex articles in the dictionary, the use of this type of comprehensive treatment could result in confusing articles for the user if it is made as an isolated lexicographic procedure. To avoid such an inconvenience for the user, the lexicographer could give a poly-accessible character¹⁰ to the dictionary by means of outer texts, particularly the back matter texts¹¹. In fact, in the proposed model certain lemmata will be regarded as salient or extremely relevant to the young Fang people, particularly lemmata concerning academic fields and cultural aspects of Fang people, who need to transfer their academic background from French to Fang, and need to know more about the culture of their parents. Most of young fang people are growing up in urban areas; they do not learn their culture in schools where everything is in French and about western culture.

The use of complex articles will give the lexicographer the opportunity to focus on specific data entries by means of outer texts procedures or article stretch internal procedures. I adhere to the use of both procedures, which is the ideal situation (cf. Gouws, 2003:67-68):

- An article internal procedure marks the specific article as belonging either to the group of complex articles or to the specific subcategorisation within the group and by doing so identifying and

¹⁰ According to Gouws (2003:65), poly-accessibility is achieved when the identification of a lexical item as lemma included in the macrostructure of the dictionary and the search route to that specific lemma sign is not exclusively dependent on the macrostructure of the dictionary. Outer texts can be utilised to establish the poly-accessibility character of a dictionary.

¹¹ It will be illogic to include additional data of complex articles in the front matter, i.e. at the beginning of the dictionary. The user would not know what that is all about.

exposing the specific data type. In the proposed model a part of the comprehensive treatment, mainly represented by cultural data, will be included at the end of each complex article.

- An outer text procedure leads to the inclusion of an outer text representing e.g. a register of the lemmata, which are the guiding elements of complex articles. In the proposed model, additional data categories related to lemmata guiding complex articles will be included in the back matter texts of the intended dictionary. That data could be complementary cultural data, grammatical data, or any type of data that the lexicographer estimate relevant for the target users. The use of complementary data in the back matter texts will allow the lexicographer to decrease the comprehensive treatment in the central list in order to avoid articles too complex and confusing for the user. The comprehensive treatment in the central list should be limited to a certain number of lines/sentences/words (it will be up to the lexicographer to decide, but should try to have a good balance about the size of the articles). Any relevant data that could take the lexicographer over the decided limit should be included in the back matter.

The use of back matter articles in the proposed dictionary model will enhance both communication-directed and knowledge-direct functions. The communication-directed function prevails because the included data could give information about the register, the morphology, etc. of the treated lemma or equivalent item, and the knowledge-directed

function prevails because the included cultural data will give a lot of information about Fang people and their culture.

4.3. Other aspects of the dictionary structure

4.3.1. Access structure

4.3.1.1. General remarks

The success of a bilingual dictionary depends largely on the success that the typical target user achieves during a normal dictionary consultation procedure. Too often these procedures fail because, although bilingual dictionaries contain all the relevant data, the structure of the dictionary and the presentation of the data is of such a nature that the target user cannot find the data or retrieve the needed information. The rapid and unimpeded access of the user to the relevant data presented in the dictionary has to be regarded as a prerequisite for successful lexicographic production in a user-driven approach (Gouws, 2000a: 102). The access structure is the search route the user follows to reach an entry in a dictionary. The search route of typical user starts with the choice of a given dictionary on account of the title of that dictionary as it is presented on the back of the cover. From there the user proceeds to the inside of the dictionary and finally reaches the article. The search route leads the user into the article to a specific microstructural entry (Gouws 2001b: 88). The procedure of accessing a dictionary and following a search route must be devised as part of the dictionary plan in the dictionary specific lexicographic process. A user-driven approach in lexicography compels the lexicographer to use all possible means at his/her disposal to

improve the quality of dictionary use and to ensure an optimal retrieval of information. Successful information retrieval can only be achieved if the data is accessible. Accessibility of data is not only determined by the use of a well-defined access structure leading the user to a venue within the word list structure of the dictionary but it is also determined by a functional positioning of the data within the word book structure of the dictionary and by employing different texts in the dictionary to explain and explicate the contents of the dictionary (Gouws, 2001a: 102). The access structure of the microstructure is called the inner access structure. All other access structures constitute the outer access structures (Hausmann and Wiegand, 1989: 329).

4.3.1.2. Outer access structure

4.3.1.2.1. Introduction

The outer access structure is the part of the search route, which leads the user from the entries on the cover of the dictionary to the lemma sign given as guiding element of the article (Gouws, 2001b: 88). The outer access structure includes all the entries on the cover (i.e. front and back cover) indicating the title of the dictionary and its typological nature as well as certain front matter texts, e.g. a table of contents, which guide the user to the relevant article. In some cases the data exposed on the back cover of the dictionary for instance immediately helps the prospective user to decide whether or not the dictionary will grant access to the information he/she is looking for. Consequently, this back cover of the dictionary constitutes a particular outer access

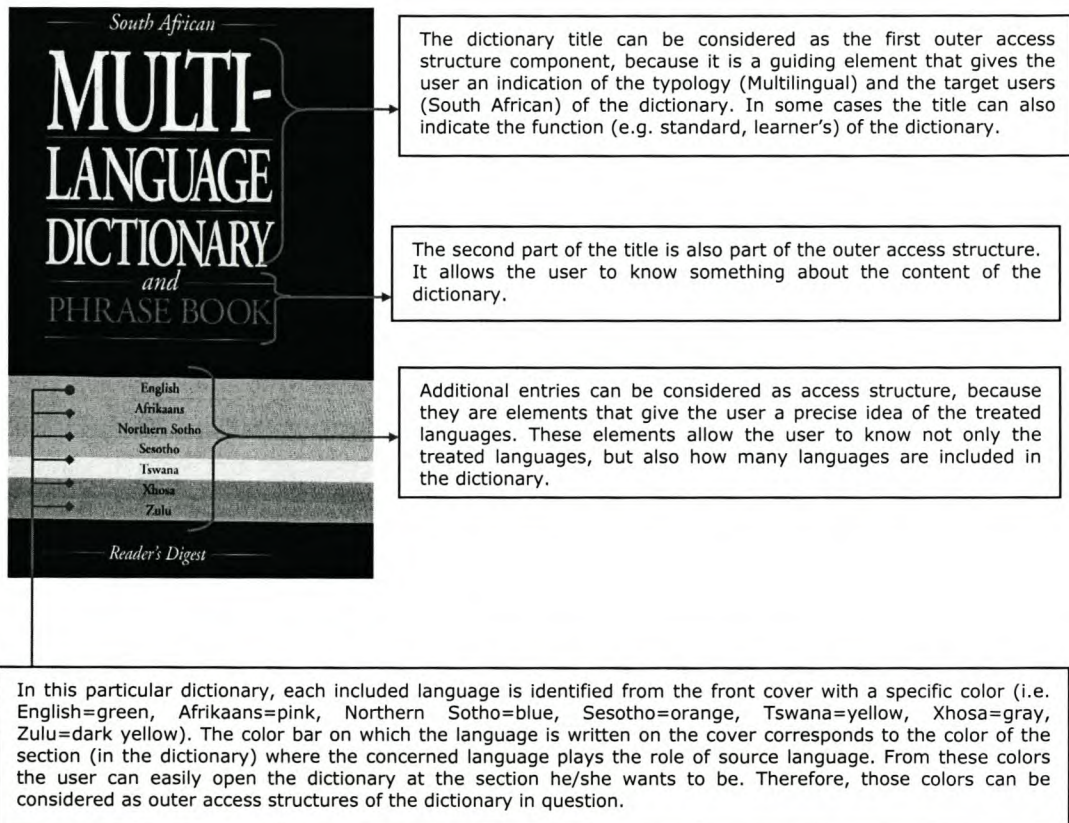
structure and **data exposure structure** (Gouws, 2002, 2002a). Most dictionaries also have the main title exposed on the spine cover.

The headers on the pages in the central list displaying key words that represent the first and last lemma sign included on a specific page also form part of the search route of the outer access structure because they assist the user in finding a specific lemma within a given article stretch. Most dictionaries have, in addition, a shorter version of the outer search path. The shorter search path starts at the running heads of the column, which are guiding elements other than the lemmata. Some dictionaries have only one running head on one page and others have two.

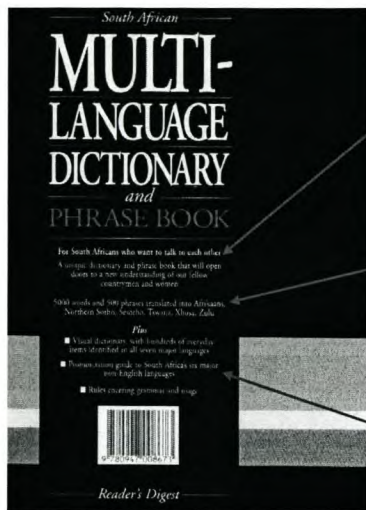
In dictionaries where I retrieved my illustration (cf. example 4), the running heads are identical to the first and last lemma on the dictionary page. In some dictionaries, there is only a partial string of the guiding element functions as a running head. On the left pages there are the first four letters of the first lemma and on the right pages there are the first four letters of the last lemma. In some cases the running heads are limited to three letters. Together the running heads of the dictionary form part of the *outer rapid access structure* (Hausmann and Wiegand, 1989: 337-338).

N.B.: Illustrations of outer access structure are shown on the following pages.

Example 1: Outer access structure on a dictionary front cover
(Dictionary cover: *South African Multilanguage Dictionary and Phrase Book*, 1996. Published by the Reader's Digest Association South Africa).



Example 2: Outer access structure on a dictionary back cover



This section on the dictionary back cover gives precise information on the number of included words (i.e. 5000) and phrases (i.e. 500).

This section of the back cover of the dictionary gives, additional information about the dictionary target users (i.e. South Africans who want to talk to each other).

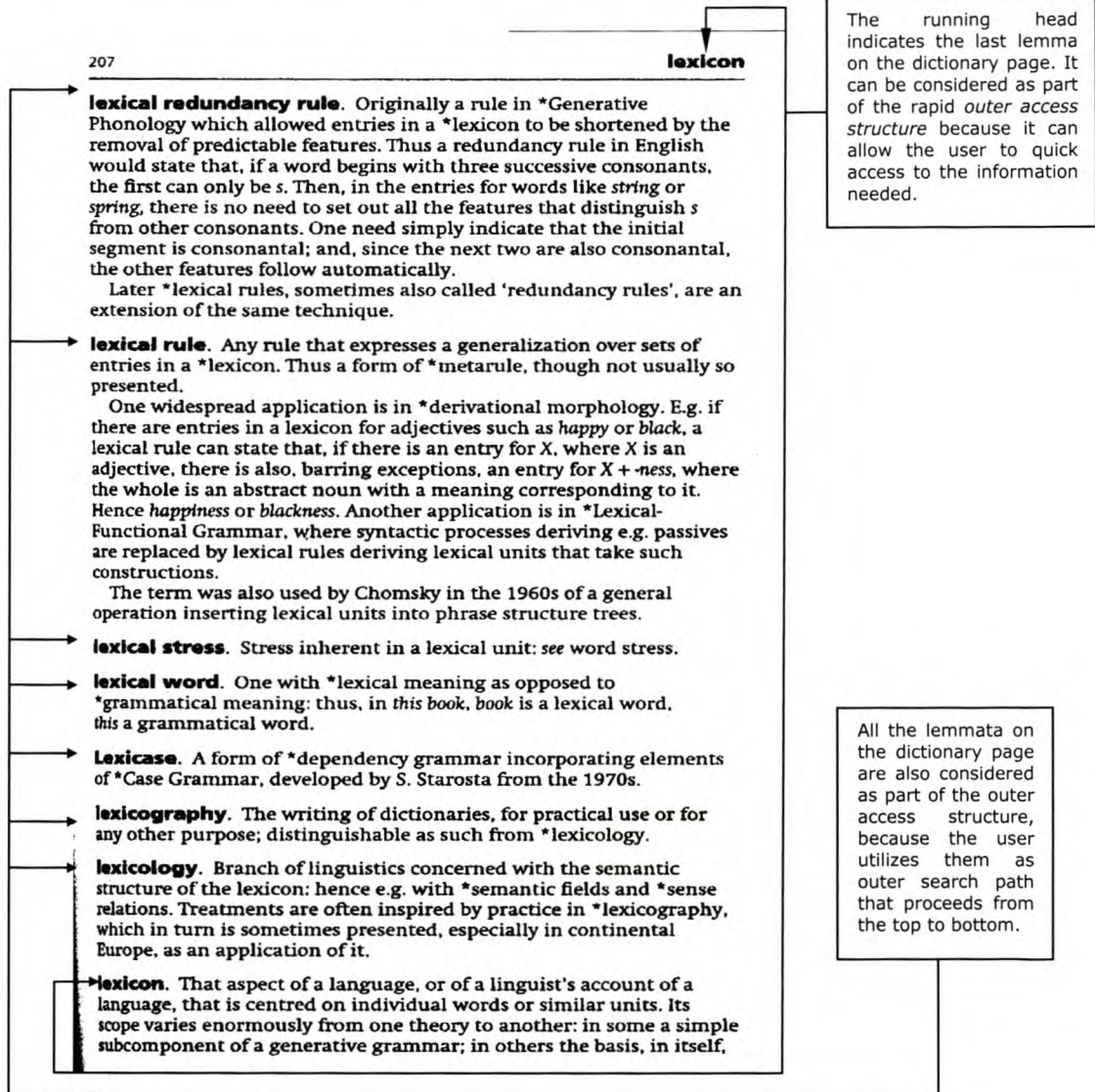
This section on the dictionary back cover gives detailed information about the dictionary contents. The number of included languages (i.e. seven), the data on pronunciation, rules in grammar and usage.

NB: The access structure here constitutes a particular outer access structure and data exposure structure.

Example 3: Outer access structure on a dictionary spine cover



The spine of the dictionary allows the potential user to identify the book he/she is looking for without opening it and without seeing the front and the back cover when the dictionary is inserted in a bookcase.

Example 4: Outer access structures on a dictionary page.

N.B. The above page is from the *Concise Dictionary of Linguistics*, (Matthews, 1997).

4.3.1.2.2. Remarks on outer access structure

As any other book, a dictionary should have its title on the front cover. However, because of the fact that it is a special book that people can consult at any time for different purposes, it is important to add additional data on the cover in order to allow users to precisely know the type of dictionary and therefore save their time. In that regard, apart from the main title on the front and back cover, a dictionary cover should also give indications about the type of data the user can find in the book. Another interesting approach that can be adopted by the dictionary compiler in terms of outer access structure is what I call the **non-typographical procedure** as part of the outer access structure. By *non-typographical procedure* I mean a method used by the dictionary compiler in order to improve the access structure of the book without using typographical markers (i.e. writing system). A non-typographical procedure can be used in many ways according to the compiler's imagination. The method used by the compilers of the *South African Multilanguage Dictionary and Phrase Book* (1996) is a good illustration in that regard. In fact, the bar colors indicating the section in which each included language is treated as source language can be considered as a non-typographical procedure as part of the outer access structure. By using this type of method the lexicographer can help the user to quickly access the dictionary section from where he/she wants to retrieve the needed information. Nevertheless, it is important to notice that the use of color bars to identify the sections of the dictionary obliges the compiler to match the relevant sections inside the dictionary with the same corresponding colors. That

procedure can be expensive and the lexicographer or the publisher may not be able to afford such an expense. Apart from the colors, other types of non-typographical procedures could also be used (e.g. pictures, icons, images, drawings, etc.).

4.3.1.2.3. Proposed outer access structure

For the dictionary model I propose, the access structure will be as functional as possible. Apart from the dictionary title (French-Fang Standard Bilingual Dictionary), the cover (i.e. front and back) will also specify the number of treated lemmata, the data categories included in the dictionary that will allow the user to know what kind of information is available in the dictionary. A rapid access structure will be available on each dictionary page. Indeed, two running heads will be presented, the first lemma at the left hand top and the last lemma at the right hand top of the page. The sections where outer texts will be included (e.g. conjugation, rules of grammar, proper names, etc.) will be identified on the dictionary cover by different color stretches.

4.3.1.3. Inner access structure

According to Hausmann and Wiegand (1989: 338), once the lemma sign is found, there is a change in the direction of the search. The *inner search path* starts at the lemma and proceeds through the dictionary article. Usually the inner access structure is explained as a string of article positions from the first to the last position. There are also *rapid inner access structures* of a linear type. They include at least one lemma, one structural indicator within the article and two search areas. Search areas are sets of article positions marked by a structural

indicator. For Gouws (2001b: 89) the inner access structure leads a user within an article to a required entry. The lemma sign is the final destination of the outer access structure and the starting point of the inner access structure. A well-devised inner access structure is characterized by the use of lexicographic conventions to assist the user on the search route. Typically, these conventions include the use of structural markers. Two kinds of structural markers are used, i.e. *typographical structural markers* and *non-typographical structural markers*.

- *Typographical structural markers* are the different typefaces, e.g. bold, italic, roman, and the use of capitals, small caps, etc. in a dictionary. The function of these markers is to indicate specific search zones or data categories. In a monolingual dictionary one often finds the lemma sign to be given in bold, the paraphrase of meaning in roman and illustrative examples in italics.
- *Non-typographical structural markers* are symbols and signs used to mark the beginning of a certain search zone or data category.

Both typographical and non-typographical markers in the inner access structure will be used for the model I propose. In fact, concerning the typographical structural markers, lemma signs in French and their translation equivalents in Fang will be presented in bold, paraphrases of meaning (in both languages) will be in normal roman characters and illustrative examples (in Fang) in italics. This presentation will allow

the user who is only interested in finding a specific category of information to have a rapid access to the article slot where the data category he/she is looking for is accommodated. As far as the non-typographical structural marker is concerned, I propose not to utilize too much specific markers, because they can confuse users. For the model I propose the following 4 symbols will be given to introduce the corresponding co-text entry:

[...] = Phonetic pronunciation

1, 2, 3...= Different senses of a lemma.

↔ = Cross-reference.

+ = Encyclopedic, cultural or ethnographical data.

It is important to specify that, to be usable or accessible to the users, the access structure of the dictionary has to be clearly explained in the front matter texts.

4.3.2. Mediostructure

According to Gouws (2001b: 91), the mediostructure is the system of reference employed in a dictionary to lead the user from one entry to another. This reference procedure may be restricted to the specific article or it may exceed the boundaries of the article. Due to the alphabetical ordering system, the macrostructural ordering in general dictionaries defy semantic relations. The lexicon of a language consists of a network of semantic relations holding between different lexical items and the application of mediostructural procedures is one of the ways to help the dictionary user become aware of the links between these items.

In the theoretical model I propose, three types of mediostructural procedures will be used:

- The first one will be a **reference marker**, i.e. a double face arrow (\Leftrightarrow) referring the user toward the treatment of other lexical item(s) where he/she can find additional data regarding the lemma or the equivalent item or any other entry in the article. The double face of the arrow means that there is a mutual relation between the treated lexical item and the one given as reference address.
- The second type of procedure will be abbreviations as *Syn.* (i.e. synonym), *Hom.* (i.e. homonym), *Ant.* (i.e. antonym), etc. referring the user to other lemma(s) or equivalent item (s) that have a semantic and/or morphological relation with the treated lemma or equivalent item. I specify that, generally in most dictionaries, the presentation of synonyms, homonyms, antonyms, etc. is not used as cross-reference. For the model I propose these abbreviations also called **reference entries** will be used as cross-reference because they will also refer to additional entries regarding the lemma or the equivalent item.
- The last type of mediostructural procedure will be the page numbers. This **complex reference address entry** involves the first two procedures. Indeed, the referred items will not be given only with the arrow or the mentioned abbreviations, but the page number indicating where they can be found in the dictionary will complete them. The page number will be given in

superscript after the referred item (e.g. *Syn.* **Nlo**²⁴⁵, ⇔ **Abo**²³). Most of the time people have to use their alphabetical knowledge to find the given reference in the dictionary. The page number procedure also called the **double address entry** will allow the user to get the referred item quicker, to access all data related to it and consequently he/she will gain time during the dictionary consultation.

In the mediostructure, two kinds of cross-references could be presented: the *article-internal cross-referencing*, which works within the boundaries of an article and the *article-external cross-referencing*, which refers a user to an entry in another article or other text in the dictionary (Gouws, 2001b: 83).

4.4. Concluding remarks

In this chapter, some aspects of the dictionary structure have been discussed. Particular attention has been paid to the data distribution structure of a dictionary, which determines the way in which data types are presented and different texts are positioned in the dictionary. Therefore, the lexicographer has to make a description of the article structure and the micro-architecture of the intended dictionary. The data distribution structure represents an ideal conjecture the lexicographer expects to apply in a practical way when compiling a dictionary. However, there is always a gap between the theory and the reality during the application. Many practical factors can make the compiler to revise or retouch some aspects of the structure of the intended dictionary. For instance during the writing

process the lexicographer may realize that some aspects are not applicable or that some specificities of the language were not taken into account during the provision. Nevertheless, in my view, the most important role of the data distribution structure is to give the dictionary maker a broad idea of where he/she is going when compiling a dictionary. The lexicographer can refer to the provisional dictionary architecture all along the compilation process and he/she should try to stay as close as possible to that plan. Thanks to the data distribution structure the lexicographer's work becomes much easier and systematic nowadays than before. For that reason it should always be an important part of the dictionary compilation process.

CHAPTER 5:

Dictionaries and language standardization

Dictionaries reflect the characteristic way in which a language has become standardized in a given situation; they influence that process by encouraging the establishment of properties and exploring new possibilities of manifestation for the functions. They can indeed become themselves the very symbol of the standard language (Gallardo, 1980:61).

5.1. Introduction

The relationship between dictionaries and language standardization that will be highlighted in this chapter particularly refers to the role dictionaries can play in the standardization process of languages, in general, and Fang in particular. In Gabonese languages, including Fang, the standardization process does not only involve the choice and acceptance of the standard dialect by the speech community, but the process also means the choice and use of a unique spelling system for the reading and writing of these languages. Indeed, no language in the world has developed without a specific and unique writing system. A language is functional through its oral and written use. However, it is through its written form that the standard form of a language can be fixed and developed. It is through its written form that a language can be taught in schools, used in administrative or official activities, etc. For that reason, in this chapter, particular attention will be given to the matter of the writing system of Gabonese languages in general and Fang in particular. The particular role dictionaries can play in that regard will strongly be emphasised. New methods will also be proposed for

the promotion and vulgarisation of the retained writing system. The main purpose of this work is to provide an efficient way for the standardization of underdeveloped languages like Fang. Therefore, apart from the dictionary, any other useful and complementary means that can help in the language standardization process should be utilized. That is why the role that new technology facilities (e.g. electronic dictionaries, Internet, etc.) could play in the development of Gabonese languages, including Fang, will be highlighted at the end of this chapter.

5.2. Standardization: a better understanding of the concept

The definition of standardization is not as easy as it seems to be. In fact, the concept of standardization has been described by many researchers; however these definitions are not always similar. While Hartmann and James (1998:131) define standardization as *a collective term for those processes, which bring about uniformity in language by reducing diversity of usage*, Matthews (1997:353) describes standardization as *the process, often in part at least deliberate, by which standard forms of a language are established*. In my point of view, standardization in the human community is **a process through which, from many ways of using or doing something (e.g. speaking a language), people will decide to recognise only one of these ways for general use in order to make the use easier and accessible to everyone, and consequently to improve the system involved**. In the specific case of a language, standardization is the process by which the choice of the standard variety is made, as well as the promotion (e.g. by teaching) of that variety in the speech community. It is important to underline that the standardization process is not

applicable only on languages. Indeed, any domain that has at least two different ways of functioning and in which harmonization is necessary in order to make the functioning identical, is using standardization (sometimes without necessary using that term). Many examples can be given in that regard. For instance, the driving system differs from one country to another. In some African countries the driving is on the right side of the road, in others it is on the left. People from countries where the driving is on the left experience many difficulties to drive in countries with driving on the right side (and vice versa), because the road signs are disposed in different ways and the driving habits are not the same.

All those differences could be the cause of many accidents and difficulties for drivers (e.g. getting a new licence when they are in a foreign country, converting their licence for local use, etc.). To avoid all these unpleasantness, one of the solutions could be the harmonization of the driving system in all African countries. Together, all African countries could decide, for instance, that from now on, the driving will be on the right. This means that all countries with left driving will adopt the right system and adapt their road signs according to that. One of the consequences will be that driving licences will be the same and will have the same value in any African country, because the driving code will be more or less the same. The choice of the driving system adopted for all African countries will be done according to specific criteria, for instance, the number of countries using the same system, the easiness of the system to be updated, the cost of the updating, etc. All these processes can be called the standardization of the driving system in Africa. Each particular driving system in each country can be called

a variety of the driving system in Africa, and the chosen system will be considered as the standard one. However, despite this common choice, each country will still retain some particularities in its driving system.

The general process of standardization can be schematised as follows:

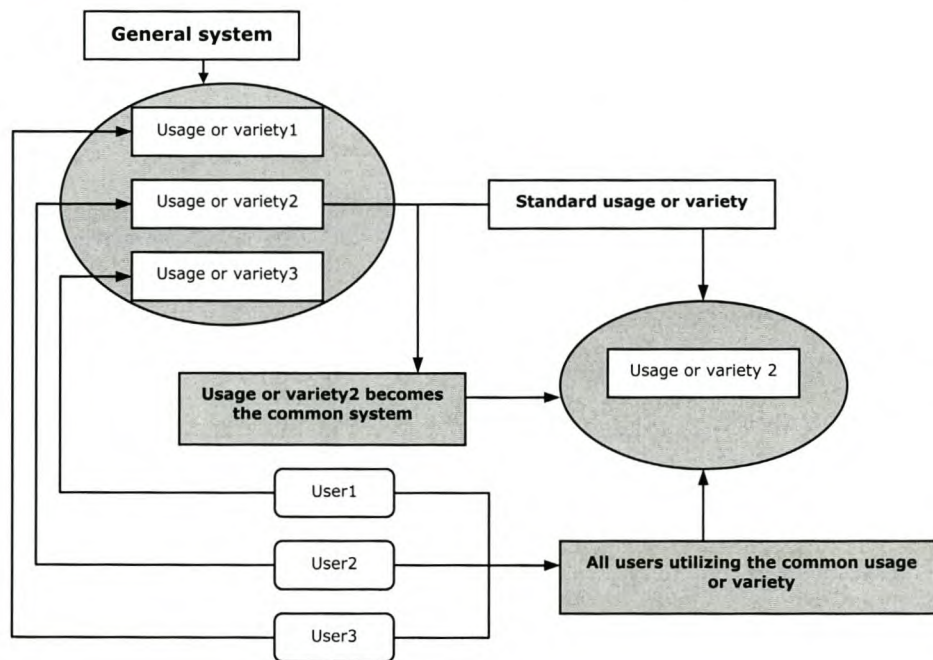


Figure 2: Global view of the standardization process

5.3. Why standardize a language?

When one asks any group of people speaking the same language if they would like to see their language develop and become a broad medium of communication in the society, the answer will obviously be unanimously positive. However, if one asks the same people if they are ready to speak a dialect different to their own one and use it in public as the standard language, the answer will be more ambiguous and hesitant. Some people will even strongly reject that

possibility. That attitude can more or less be explained by the fact that most people do not know that the development of a language means the harmonization of the latter, i.e. the reduction of varieties in order to allow a better mutual understanding between speakers. The standardization of a language is necessary to facilitate communication in a multilingual or multidialectal community. It also serves to establish an agreed orthography and provide a uniform form for schoolbooks. This will not only determine the form taught in language classrooms to native as well as non-native speakers, it will also decide the language of learning generally (Anthonissen & Gough, 1998:40).

As any organized system, the standardization of a language should raise the improvement or development of the latter. It is easier to develop a system when it is standardized than when it is not. That can be seen in the fact that nowadays most products have many common features, even when they are from different manufacturers. For instance many products have been standardized in the technological field in order to ease their accessibility and use. Computers are a very good illustration in this regard. Indeed, most software from different manufacturers (e.g. Microsoft, Netscape, etc.) can be used in most computers from different makers (e.g. IBM, Compaq, Macintosh, etc). The functioning of computers has been standardized in such a way that the user can also update and/or repair his/her machine without too much difficulty. Because of the standardization of computers, the compilation of software and computer programmes has developed exponentially. As with computers, the standardization of a language is a part of its development process. "The suppression of variation in language will

ensure communication over longer distances of space and time with minimum of misunderstanding” (Anthony Lodge, 1993:24).

5.4. Standardization and the writing system

5.4.1. Introduction

Gabonese languages are still searching for their real place within the modern linguistic concert. Indeed, these languages still need to be exhaustively identified, classified and described in order to allow their preservation and their development. It is attempted to use languages actively in all fields of the modern life where their speakers carry on their activities. However, to achieve that objective, Gabonese languages must necessarily go through the same process used by the so-called modern or international languages (i.e. English, French, Spanish, etc). According to my opinion, this process starts with the approval of a unique writing or alphabetical system accepted and used by everyone. A standard alphabet without which any attempt of developing Gabonese languages would remain vain. Therefore, the current debate (i.e. in the Gabonese linguistic field) on the development of Gabonese languages cannot avoid the crucial problem of the alphabet. “Standardization is more easily achieved in writing than in the speech” (Anthony Lodge, 1993:24).

5.4.2. History of the alphabet¹

5.4.2.1. Introduction

An alphabet (from *alpha* and *beta*, the first two letters of the Greek alphabet), can be defined as a set of written symbols, each

¹The essential part of the information presented in this section has been retrieved from the website <http://histoire.typographie.org/carateres>.

representing a given sound or sounds, which can be combined to form all the words of a language. An alphabet attempts ideally to indicate each separate sound by a separate symbol, although this end is seldom attained, except in the Korean alphabet (the most perfect phonetic system known) and, to a lesser degree, in the Japanese syllabaries. Alphabets are distinguished from syllabaries and from pictographic and ideographic systems. A syllabary represents each separate syllable (usually a sequence of from one to four spoken sounds pronounced as an uninterrupted unit) by a single symbol. A pictographic system represents picturable objects, for example, a drawing of the sun stands for the spoken word *sun*. An ideographic system combines various pictographs for the purpose of indicating non-picturable ideas. Early systems of writing were pictographic-ideographic, e.g. the cuneiform of the ancient Babylonians and Assyrians, the Egyptian hieroglyphs, the written symbols still used in the Chinese and Japanese languages, and Mayan picture writing. What converts such systems into an alphabet or syllabary is the use of a pictograph or ideograph to represent a sound rather than an object or an idea.

The history of the world has shown that it has always been judicious for human societies to get inspired from the past or from the others in order to solve problems with which one is confronted, and the development of languages does not escape that reality. To succeed in the development process of its languages, Gabon should use the example of advanced countries in the linguistic field. As far as the orthographical alphabet is concerned, such a step implies a brief historical overview of the development of alphabets in countries where the development of languages is proven and

recognized by all. This overview will enable us to come up with some suggestions, which could be helpful in the choice and the stabilization of the alphabetical system of the Gabonese languages.

5.4.2.2. Origins of the alphabet

The first writing systems were born in various forms at different moments in many places over the world (e.g. Mesopotamia, Egypt, China, Central America, etc). However, it seems that the birth of the first alphabetical writing system is geographically localized in Canaan, around the second millenium BC, in the area that corresponds to the actual Middle East (i.e. Lebanon, Israel, Syria, Jordan and the Sinai). The ancient Semitic alphabet originated from the Egyptian civilization. This pseudo-hieroglyphic writing was functioning according to the principle of the acrophony, i.e. each pictogram symbolized the very first sound of the word represented. In Semitic languages, every word starts with a consonant. Consequently, the pseudo-hieroglyphic alphabet was essentially consonantal. At the same time around the 14th century in Ugarit², on the Phoenician coast, a consonantal alphabetical writing of 30 signs was invented using the wedge-shaped graphic system utilized in the ancient Acadia³. Before the end of the 12th century BC, the traditional alphabet of 22 letters was established after one millennium of evolution since the invention of the hieroglyphs. The written form was stabilized just as the direction of the reading, which was done henceforth from the right-hand side to the left. The

² Ugarit (also known as Ras Sharma), capital city of a Late Bronze Age principality of coastal northern Syria. It was first settled around 6500 BC, during the Neolithic period, and was inhabited until 1180 BC (the end of the Late Bronze Age). During the Middle Bronze Age (2000-1600 BC), Ugarit was already an important site and had strong trading connections with Egypt (Microsoft Encarta, 2001).

³ Original name of the parts of Canada now divided between the Maritime Provinces of Nova Scotia, New Brunswick, and Prince Edward Island (Microsoft Encarta, 2001).

significant acquisition at that time remained the use of a reduced set of graphic signs to symbolize the articulated language.

5.4.2.3. Birth of vowels

The use of the Phoenician alphabetical writing encountered some difficulties in the writing of Greek and the majority of the Indo-European languages (e.g. Persian, Sanskrit). The main difficulty related to any syllabic writing is to isolate a consonant, when the latter is not followed by a vowel. However, groups of two or three consonants are current in Greek, and consequently a Greek text whose vowels are not noted is thus completely unreadable. Pragmatic Greeks transformed the Phoenician alphabet and adapted it according to their language. Nevertheless, the most significant invention of the Greeks is the allocation of a value they did not have before to certain Phoenician letters, i.e. the value of vowel. From there were born alpha (**A**), epsilon (**E**), omicron (**O**), and upsilon (**Y**). For **I** they invented from nothing a letter, i.e. the iota. This "light of vowels" to paraphrase Etiemble, is the decisive contribution of Greeks in the history of our civilization.

5.4.2.4. The wider international influence of the Greek alphabet

The Greek alphabet inspired many civilizations. Etruscans whose civilization appeared in the current Tuscany at the 7th century BC also took the Greek alphabet to transcribe their language. Etruscan Kings reigned in Rome up to the 4th century BC, when Latin people chased them out. These Latin, i.e. the future Romans, adopted the Etruscan alphabet to transcribe their language. Therefore, around the 3rd century BC, an alphabet of nineteen letters was established,

after the reintroduction of X, Y and Z in the alphabet at the 1st century BC during Cicerone epoch. It is that alphabet, known today as Latin's alphabet, that the majority of Latin civilizations (e.g. France) had adopted.

The following figures are some examples of the development of writing in some civilizations.

| | | | | | | |
|---------|--|--|--|--|--|--|
| ox | | | | | | |
| plough | | | | | | |
| orchard | | | | | | |
| sun | | | | | | |
| grain | | | | | | |

Figure 3: Development of the Sumerian into cuneiform (Source: Ndinga-koumba-Binza, unpublished)

| HIEROGLYPHIC | | | HIERATIC | | | DEMETIC |
|--------------|--|--|----------|--|--|---------|
| | | | | | | |
| | | | | | | |

Figure 4: Development of the Egyptian writing (Source: Ndinga-Koumba-Binza, unpublished)

| | Egyptian Hieroglyphic | Sini | West Semitic | Early Greek | Late Greek | Latin |
|----------------|--------------------------|------|-----------------|----------------|---------------|-------|
| ox | | | | | A | A |
| head | | | | | P | R |
| snake | | | | | N | N |
| mount- ains | | | | | Σ | S |
| court- yard | | | | | B | B |

Figure 5: Development of five letters from the Egyptian hieroglyphs (Source: Ndinga-Koumba-Binza, unpublished)

5.4.3. The writing of African languages

5.4.3.1. A brief overview

The role of the Egyptian hieroglyphs in the development of writing around the world shows that writing is not new in Africa. The sub-Saharan Africa has also made its contribution to the evolution of writing during the colonial era. Many evidences such as the rock drawings, found in many regions of the continent or the graphic symbols of the Komo society among the Bamanas, still confirm that (Touré, 1990: 56). However, these writing systems did not really develop because of the early advent of the Arabian and Western civilization in the continent.

Missionaries and colonial administrators, with the objective to evangelise and communicate in local languages, were the first ones to bring Latin writing systems into African languages. Indeed, dictionaries, grammars, glossaries, schoolbooks, lexica, etc. compiled by them were written in the Latin alphabet. The Xhosa provides a good illustration in this regard. According to McLaren's grammar (1936), the same alphabet is used for the writing of English and Xhosa, except for the consonantal sounds *b*, *r*, and *ǀ*. Xhosa alphabet, at that time, had 29 symbols including five vowels (*a*, *e*, *i*, *o*, *u*) and three clicks (*c*, *q*, *x*). Clicks do not exist in English; however they have been transcribed by three symbols from the English alphabet. The situation happened in many African countries, where African languages have been written with Western writing systems. After their independence, most African countries really started to think about adapting their writing systems. The outcome of this awareness had been the organization through the continent of national and international meetings in order to equip

African languages with local or regional writing systems reflecting their specific needs. From these consultations were born most of the current writing systems in the majority of African countries.

5.4.3.2. Alphabet and orthography

5.4.3.2.1. Definitions

Before any discussion regarding the matter of the alphabet and orthography in Gabonese languages, a brief analysis of the two terms is important in order to see if they are related or are two completely different concepts. The term alphabet results from the combination of *alpha* and *beta*, which are the first two letters of the Greek alphabet, and which would correspond to the letters A and B in the English alphabet. Nowadays, this term is used in many languages to indicate the same reality. The *Nouveau petit Larousse* (1971) described the alphabet as a *list of all the letters from a language*⁴. The *Dictionnaire Universel* (1995, 2nd edition) supplements this definition by describing the alphabet as the *set of letters being used to transcribe the sounds in a language*⁵. *Cambridge International Dictionary of English* (1995) does not give a definition very different from the first two, it describes the alphabet as *a set of letters arranged in a fixed order, which is used for writing a language*.

Microsoft Encarta Encyclopaedia Standard (2001) is more precise. It describes the alphabet as *a set of written symbols, each representing a given sound or sounds, which can be combined to form all the words of a language. An alphabet attempts ideally to*

⁴ Translated from French

⁵ Translated from French

indicate each separate sound by a separate symbol. All the definitions of the term alphabet, whatever the language, are summarized with the fact that it is a set of letters or symbols being used for the writing of a given language.

When the alphabet is described as “a set of letters or symbols being used for the writing of a given language”, the orthography or spelling will be defined as “the conventionalised system of representing speech by writing in a particular language (Hartmann & James, 1998:130).

It is necessary to keep in mind that the notion of alphabet can encompass two distinct aspects of language transcription: the phonetic aspect (i.e. phonetic alphabet) and the orthographical aspect (i.e. orthographical alphabet). One speaks of the orthographical alphabet when the letters or symbols are used for the writing of words in a language. Therefore, the alphabet is not to be confused with the orthography or spelling. The orthography or spelling is *a regulated system in which letters from the alphabet are combined to form words (and consequently sentences) in a language*. Thus, the orthography can be compared to an organized system in which the alphabet serves as tool. One speaks of orthographical alphabet to indicate letters or symbols, which are used to form words in a language according to the specific orthographical rules.

The situation is not the same with regard to the phonetic alphabet. Indeed, phonetics deals with the sounds of the language. It makes possible the transcription of sounds of the language exactly in the way they are pronounced. The phonetic alphabet first aims to allow the transcription of sounds in a language, even when

the user does not know the orthographical alphabet and/or the orthography of the language.

As far as Gabon is concerned, the matter concerning the writing of Gabonese languages does not really concern the phonetic alphabet, because there already exist phonetic alphabets such as IPA, IAI and ASG (i.e. Scientific Alphabet of Gabonese Languages), that are appropriate enough to be used for the transcription of sounds in these languages. The problem that Gabonese languages encounter concerns the orthographical alphabet. Actually, the matter is not about the quality of the existing alphabets, but it is about the necessity of choosing only one of them as the standard one that should be used by everyone. Therefore, the real question is: what orthographical alphabet should be used for the writing of Gabonese languages (including Fang)?

5.4.4. Writing of Gabonese languages

5.4.4.1. Introduction

As most civilizations, the Gabonese population was using some graphical means of communication long before colonization. Many rock drawings were found in some parts of the country (e.g. the prehistoric cave of Bongolo). However, the modern form of the Latin writing was introduced in Gabon with the arrival of missionaries (i.e. Catholics and Protestants) and colonial administrators. It was employed for the translation of religious documents (e.g. the Bible, catechism, hymn, etc.), as well as the transcription of local names (e.g. people names, villages, ethnics, etc.). Despite all these efforts, Gabonese languages were still considered as oral languages because of the diversity of the writing systems used and the fact that these systems were derived from foreign languages

(particularly French). Pierre Alexandre (as cited by Kidda Awak, 1990: 11) estimates for instance that Fang got “three Catholic and two Protestant transcriptions for more or less one million of speakers”. Mayer (1990) talks about three main periods in the development of the writing in Gabon: the exploratory, the pedagogical and the scientific writing.

- The first period (i.e. the exploratory) corresponds to the arrival of the first Europeans on the Gabonese coast. That period also corresponds to the arrival of explorers (mainly Portuguese) around the beginning of the 16th and the mid 19th century. At that period the writing system was entirely based on western ones. Mayer (1990: 67) specifies that this period, which corresponds to the oldest writing of our languages, has furnished Gabonese languages three centuries of writing trails. These trails are essentially visible on the maps, which carry names of people and areas.
- The second period concerns the so-called pedagogical period of writing in Gabon. This phase, starting in 1850, opened the learning and inclusion of Gabonese languages into schools, mostly by missionaries. However, the alphabet and the writing systems used at that time were still largely based on European ones (Mayer 1990: 73).
- The third and last period (i.e. the scientific) established by Mayer (1990: 88-89) starts in the second half of the 20th century. It is characterized by the consideration and the use of phonetic and phonological specificities of Gabonese languages for the establishment of new writing systems. The

current period is more or less the continuation of the third phase, because scientific works, discussions and debates concerning the writing of Gabonese languages nowadays are still taking place.

The purpose of this section is to present some of the systems proposed for the writing of Gabonese languages. Among all of them, only three will retain my attention: the *Raponda Walker's alphabet*, the *April 1999' alphabet*, and the *Rapidolangue's alphabet*. The presentation and analysis of these alphabets will allow me to propose some ways that can be useful in the choice of the standard alphabet for the writing of Gabonese languages.

5.4.4.2. The Raponda Walker's alphabet

In the article entitled *Alphabet des idiomes gabonais*⁶ (re-edited in 1998 by la Raponda Walker Foundation) published for the first time in 1932 in the *Journal de la Société des Africanistes*⁷, Raponda-Walker established a list of phonemes with corresponding graphemes by using the IPA and Africa Alphabet. He proposed a system of more or less 36 alphabetical symbols. That alphabet includes 28 consonants and 8 vowels, and is compiled according to the principle of aphorism, i.e. to each grapheme corresponds only one unique symbol to transcribe it, as one can see in the following table.

⁶ Alphabet of Gabonese idioms

⁷ Journal of the Africanist societies

| Vowels | Simple Consonants | | | Consonants with diacritics |
|--------|-------------------|---|---|----------------------------|
| a | b | l | v | ĉ |
| e | c | m | w | |
| è | d | n | x | ĝ |
| i | f | p | y | ñ |
| o | g | q | z | ṇ |
| ö | h | r | | ṛ |
| u | j | s | | ẉ |
| ü | k | t | | |

Table 6: Alphabet of Gabonese Idioms from Raponda Walker

According to Raponda-Walker (1998: 7), this orthographical alphabet, thanks to the specific diacritic signs used with different letters, easily reproduces the pronunciation of native words. One can also remark that no prosodic phenomenon is mentioned. Raponda Walker's alphabet is very close to the Latin one (Idiata, 2002: 50), as one can observe from the symbols c, k and q that represent the sound [k].

5.4.4.3. The April 1999 alphabet

Ten years after the creation of the Scientific Alphabet of Gabonese languages (ASG), the Ministry of Education, with the support of UNESCO, organized on April the 8th, 9th and 10th 1999 a consultation of experts (i.e. linguists, sociologists, anthropologists, etc.) with the purpose to come up with a set of symbols or graphemes for the writing of the Gabonese languages. From that consultation was born the April 1999 alphabet, a set of 35 graphemes, i.e. 26 consonants and 9 vowels. The experts also recommended the transcription of five tones for the writing of Gabonese languages. According to the final report of that workshop, the choice of these graphemes was based on their simplicity and functionality.

| Vowels Signs | Tones Denominations | Consonants | | |
|-----------------|------------------------|------------|----------|----|
| | | Monographs | Digraphs | |
| a | high tone | b | m | gh |
| e | low tone | c | n | jh |
| ɛ | medium tone | d | ɲ | sh |
| ə | downward tone | ɗ | p | vh |
| i | rising tone | f | r | ny |
| o | | g | s | |
| ɔ | | h | t | |
| u | | j | v | |
| ɯ | | k | w | |
| | | l | y | |
| | | | z | |

Table 7: The April 1999 alphabet

The April 1999 alphabet, also called the *Nouvel Alphabet des Langues du Gabon*⁸ (ALG), has the advantage to be born from an initiative of the Ministry of Education and it is an outcome of a workshop of experts. Therefore, it is the result of broad thinking, and consequently deserves a better consideration than what has been the case until then. The only reproach that can be made about that proposition is the symbol ∂ , which is identical to the phonetic symbol. In my view, it is important to make a distinction between orthographical and phonetic symbols, particularly for the vowels. That distinction will allow anyone who wants to learn the language (i.e. the pronunciation) to use phonetics as bridge. To solve that matter I propose the utilization of an underlined **a** (**a**) instead of ∂ for the orthographical representation of the latter.

5.4.4.4. The Rapidolangue's alphabet

The third and last system that could play the role of standard orthographical alphabet for Gabonese languages is the one used by

⁸ New Alphabet of Gabonese Languages

the Raponda Walker Foundation in the *Rapidolangu* books⁹. Symbols proposed in that alphabet are a synthesis of the Raponda-Walker and April 1999 alphabets. Because that alphabet is the most recent, it has the advantage of taking the deficiencies of the previous ones into account.

| Consonants | | Vowels | |
|---------------------|----------|-------------|---|
| b | m | a | i |
| d | n | ä or â | o |
| f | s | e | ö |
| g | t | ë | u |
| k | z | è | ü |
| l | | | |
| Combined consonants | | Semi-vowels | |
| ng | ny or gn | y | w |
| gh or h | dy or dj | | |
| ŋg | dz | | |
| mb | mp | | |
| | nd | | |

Table 8: The *Rapidolangu*' alphabet

The *Rapidolangu*'s alphabet is a set of 24 consonants, and 11 vowels. That alphabet is very close to the Latin one and that gives it a certain accessibility to the laypersons in Gabon who are more accustomed to the French alphabet.

5.4.4.5. Quality of the standard alphabet

Many Gabonese researchers (particularly linguists) estimate that the majority of writing systems used by missionaries and colonial administrators for the transcription of Gabonese languages are too similar to the French alphabet, and they think that symbols from these alphabets do not accurately take into account the realities of these languages. In my opinion, this vision is a bit restricted. Before

⁹ The *Rapidolangu* books are a set of multilingual books presenting French lexical items and their translation equivalents in English and some Gabonese languages (e.g. Fang, Miénè, Yipunu, Yinzébi, Lembaama, etc.)

taking such a position, it is necessary to highlight the following aspects:

- For historical reasons, the French alphabet is the writing system with which the Gabonese population is most accustomed. It is the system used by everyone to read and write in all official activities (i.e. school, work, media, etc.). For that reason, it is the system that is better known and generally used in Gabon, because of the status of French as sole official language in this country. Therefore, it is not logical to speak about the matter of the writing of languages in Gabon without considering this aspect.
- The first works (i.e. in lexicography and linguistics) that have been compiled in Gabonese languages have used the French alphabet. However, the experience shows that these works had a more significant impact on the reading and writing of these languages, compared to those compiled nowadays according to scientific system criteria. As an illustration one can easily recognize that religious texts (e.g. Bibles) and other books (e.g. dictionaries, lexica, glossaries, spelling-books, etc) compiled in Gabonese languages with the French alphabet have allowed many people to be able to read and write easily in Fang, Yipunu, Omiénè, Yinzébi, etc. Today many old Gabonese people can still read and write in their languages thanks to the French alphabet (particularly in religious fields).
- As shown in the historical overview, the majority of alphabets that exist nowadays in many languages of the world

(especially the Indo-European languages) are actually the result of secular transformations of Egyptian and/or Greek writings. Each country, culture, civilization or people have adapted symbols according to their needs and context. That situation happened in some African languages that are advanced enough in their writing systems. Swahili for instance, one of the most developed African languages, was already a written language in Arabic script in the 13th century (Eastman, 1983:21 as cited by Mdee, 1999:120). This means that it is not the origin of the symbols that is important, but their use or adaptation according to the language.

- Finally, one should not forget that any alphabetical system is improvable and adaptable. Therefore, it does not have to be perfect from the beginning. It can be improved progressively and it is only with use that one detects its insufficiencies. For instance, one can observe the absence of symbols to transcribe certain sounds or the difficulty of bringing the transcription closer to the pronunciation.

These remarks enable me to make the following deductions:

- There are no particular criteria for the choice of alphabetical symbols. It is a convention, and one can decide to use any kind of representation (e.g. drawings, images, icons, signs, graphs, etc.) as alphabetical symbols. The only condition is that each sound of the language must correspond with a particular symbol in the alphabet. In addition to that, it is necessary to make sure that all relevant or pertinent sounds of the language are definitely represented.

-
- What makes an alphabetical system strong is its acceptance, its teaching, its use and especially its adaptation. All languages with a long written tradition (e.g. French, English, Afrikaans, Spanish, German, Greek, etc.) took hundreds, even thousands of years before they reached the current final versions of their alphabetical systems. Therefore, it is utopian to think that any consultation of experts, any seminar or workshop will generate an orthographical alphabet that is immediately perfect and compatible with all the Gabonese languages. All past, current or future proposals of the alphabets should be open for changes in order to improve them.
 - The choice or the creation of the alphabet, in the majority of the so-called developed languages, has always been made without ignoring what already existed. The former Semitic alphabet, for example, is a continuation of the Egyptian civilization, whereas the Greeks transformed the Phoenician alphabet and adapted it according to their language. Therefore, the choice of an orthographical alphabet based on the French one is not a "crime" against Gabonese languages. The most important thing is to readapt that alphabet according to the particularities of these languages. On the contrary, the familiarization of Gabonese people with the French orthographical system can be a good thing if the selected alphabet is close to French. The population will learn and understand it more easily.

5.4.4.6. Choice of the standard alphabet

First of all, it is important to note that if it is scientifically possible to obtain a unified orthographical alphabet for all the Gabonese languages, it is on the other hand impossible to have one orthographical system for all these languages. Indeed, one should remember that the **orthographical alphabet** is a *set of letters or symbols used for the writing of a language*, whereas the **orthographical system** is a *set of rules governing the functioning of each language*. Consequently, since all Gabonese languages do not function in the same way; they cannot have the same orthographical system. Therefore, speaking of one orthographical system of Gabonese languages is not logical. However, one can speak of the orthography of Fang, Yipunu, Yinzébi, Téké, etc.

In the majority of works related to Gabonese languages the term orthography is used, not to refer to the rules governing the writing of words, but to speak of the symbols used for the transcription of words in these languages. Seen in this way, the expression orthography of Gabonese languages can be understandable. In this work, to avoid any confusion, I will only use *orthographical alphabet* when I am talking about the orthographical symbols and *phonetic alphabet* when I speak of the phonetic symbols, and *orthographical system* when speaking of rules guiding the functioning of the language.

As previously said, up to now, several orthographical alphabets have been used for the transcription of Gabonese languages. This situation does not favour the development of these languages. The development of any language is difficult and slow without a unique writing system. To make a significant step in the

development process of Gabonese languages, the matter of the orthographical alphabet has to be solved as quickly as possible. A consensus regarding the use of one unique writing system has to be reached.

Generally, the standardization of a language is based on two main steps: (1) the creation of a model for imitation, and (2) the promotion of this model over rival models (Ray as cited by Mdee, 1999:120). The language standardization process inevitably obliges the choice of one of the language forms as standard one. This form will be used as reference for any official and academic use of the language. However, to promote the standard form, particularly in the academic field, it must have a written form. The standard form of a language stabilizes and develops not orally, but in the written form. That is why it is necessary to have a standard alphabet, which will allow the promotion of the standard model. Therefore, the standardization of Gabonese languages should go together with the standardization of the alphabet.

The fact that Gabon, compared to other African countries, is behind in the development process of its languages can be considered somehow as an advantage. Indeed, instead of wasting time and money in order to seek an ideal way to achieve the development of its languages, Gabon could/should get inspired by the examples of countries already advanced in that domain. In the majority of countries, the choice of a standard dialect or alphabet was often the result of the intellectual or political influence.

In France, two main steps characterized the history of the development of French in the middle Ages. The first step took place with the active influence of the social and cultural elite. They made

use of standard French on all administrative, diplomatic and literary levels. The second step, still under the influence of this same social and intellectual elite, was made with the marginalization of the other French dialects that were restricted to be used only in popular fields, particularly in rural areas (*Dictionnaire Universel*, 1995). In Africa, several countries such as Ivory Coast (National Alphabet of Ivory Coast), Nigeria (Pan Nigerian), Guinea, etc. worked out national alphabets adapted to the transcription of all their languages. The adoption of these alphabets was often made by means of political decisions, as it was the case in many countries in Africa e.g. Guinea (Touré, 1990).

It could be necessary, even indispensable that the Gabonese Government intervenes in the process of the choice of the standard dialects and alphabets. Its intervention will make things easier and quicker. If the chosen alphabet takes into account all the Gabonese languages, as well as their varieties, it could be easily used for the orthographical transcription of any standard dialect. Therefore, the choice of the orthographical alphabet will be very helpful for the standardization of Gabonese languages. Seeking consensus between linguistics experts (i.e. linguists, lexicographers, etc.) is of course important for the choice of the standard orthographical alphabet, but when that consensus becomes difficult to be reached, the Government intervention has to take place after consultation with experts. It is, in my view what should be done in the Gabonese context. It is necessary to underline the fact that the Government decision in that regard will not automatically and unanimously be accepted at the beginning, but at the end everyone will stick to it. The experience has proven that things are going faster when they

are under the influence of the Government than when one leaves the people to decide through the time.

One must keep in mind that choosing an alphabet does not automatically resolve the matter of orthographical transcription in a language. There is no perfect system. That means that the chosen system is open to changes, improvements and adaptation. It is for instance possible to realize later on that some sounds in some languages cannot be written with the chosen alphabetical symbols. In that case, new symbols can be added or the existing ones can be adapted in order to allow the transcription of all the sounds of the language. If one wants to know whether the chosen alphabet is good or not, it has to be used and experimented in order to detect its weaknesses and be improved according to that. Therefore, the sooner the alphabet will be used the sooner it can be improved. The research of an ideal alphabet can take several years, even decades, particularly in the case of a country like Gabon where the inventory of local languages is not exhaustive yet.

5.4.4.7. Matter of tones in the writing of Gabonese languages

5.4.4.7.1. Introduction

It is difficult to talk about the writing of Gabonese languages without taking into account the matter of tones. Indeed, the issue of the writing of Gabonese languages is not only related to the choice of the alphabetical system, but also to the dilemma of the transcription or not of the supra-segments. In linguistics, a tone is a phonetic or phonological unit belonging to a set distinguished or primarily distinguished by levels of or by changes in pitch (Matthews, 1997:379), e.g., in Fang *àbò* "foot" and *àbó* "to do" are distinguished phonologically by a low tone (ò) and a high tone (ó)

on the segment (o). As most Bantu languages, Gabonese languages are tone languages, i.e. languages in which each syllable is characterized phonologically by a distinct tone or sequence of tones. In tone languages, contrasts of pitch have some lexical or morphological role. The purpose of this section is to provide some ways in order to solve the matter of tones in the writing of Gabonese languages. What types of tones can one find in Gabonese languages? Should one note or not note these tones? The aim of this section is to provide answers to these questions. It is also hoped that these answers will be of assistance in the improvement of the orthographical writing of Gabonese languages, in general, and Fang in particular.

5.4.4.7.2. Types of tones in Gabonese languages

Most tones in Gabonese languages have significant distinctive functions, in the phonological system as well as in the vocabulary. Generally, tones are grouped into two main categories: punctual tones and modulated tones. A punctual tone is the one which, systematically fixed on one syllabic unit, is identified by only one aspect of the melodic pitch (i.e. high or low). A modulated tone is identified by successive changes of the melodic pitch during the pronunciation of a syllable. Punctual tones are more regular in Gabonese languages.

Different linguistic descriptions in Gabonese languages up to now have identified three categories of tones, i.e. main, intermediary, and secondary tones (Afane Otsaga, 2002:77).

- Main tones, more regular and widely spread in Gabonese languages, are essentially high and low punctual tones, as well as rising and downward modulated tones.
- Intermediary tones have a lesser presence and spreading than the main ones in Gabonese languages. They are recurrent enough and easily identifiable. Among the intermediary tones, one has medium, high downward and supra high tones.
- Secondary tones have a lesser presence and spreading than the first two categories. They can be found in few languages. Three tones belong to that group: infra low, very high lowed and very high-downward.

In deduction, nine types of tones can be found in Gabonese languages. The following table gives some examples of words with different types of tones in some Gabonese languages.

Table 9: Examples of words with tones in some Gabonese languages

| | Tones | Examples | Meaning in English | Language |
|---|--------------------|----------|--------------------|----------|
| 1 | low | sè | <i>yard</i> | Jake |
| 2 | medium | íyǎ̀lè | <i>power</i> | Benga |
| 3 | high | dóli | <i>money</i> | Yipunu |
| 4 | downward | mbó | <i>arm</i> | Sake |
| 5 | rising | díkǎ | <i>spear</i> | Seki |
| 6 | very high | ngǔ̀yì | <i>boar</i> | Gisira |
| 7 | infra low | wùlú | <i>foot</i> | Lendumu |
| 8 | very high lowed | múti | <i>tree</i> | Liduma |
| 9 | very high downward | ngáânzi | <i>roots</i> | Yipunu |

5.4.4.7.3. Pros and cons of transcription of tones

Two points of view are dividing the scientific community working on the Gabonese languages. The first point of view estimates that the transcription of tones in the writing of Gabonese languages is essential even obligatory, since the latter play a relevant role in the

functioning of these languages. The supporters of this point of view often argue that as the accents are noted in French, tones must be noted in Gabonese languages. The second point of view is that of those who think that the transcription of tones is not useful for the users, particularly laypersons in linguistics. Moreover, the good reading of tones (i.e. their pronunciation) is accessible for people aware of linguistic matters and not to common people. The following remarks can be made as far as the first point of view is concerned:

- It is not wise to compare the transcription of accents in French with that of tones in Gabonese languages. The two types of supra-segments (i.e. accents and tones) do not function in the same way in both languages. In French, accents are integral parts of the alphabetical symbols (e.g. é, è, ê). These symbols are always accompanied by their supra-segmental elements whatever their situation or their context is (i.e. isolated or in a sentence). Tones in Gabonese languages (as in all tone languages), are not linked to a particular or specific segment. They can appear above any type of syllabic segment according to the melodic curve of the word. Therefore, it is difficult to associate a particular segment to a particular supra-segment. In deduction, the comparison between accents in French and tones in Gabonese languages is not relevant.

Kwenzi-Mikala (1988: 209), one of the researchers advising the transcription of tones in Gabonese languages, estimates that if one wants to give Gabonese languages a written form readable by everyone, it is essential to note tones. According to Idiata (2002:

61), the matter of tones in Gabonese languages is not about the fact that they have to be noted or not. For him tones have to be transcribed, because they are playing a role as important as consonants and vowels. The real matter for Idiata is to find out the best economical and functional way tones can be transcribed. Kwenzi-Mikala (1988: 209), suggests in that regard the transcription of only one tone, i.e. the most frequent tone in the concerned language.

Idiata (2002: 62) proposes that the matter of tones should be treated differently according to the particularities of each language, because some languages are more complex than others. They do not function in the same way. Therefore, each language should be equipped with a particular system as far as the transcription of tones is concerned. In the same vein, the consultation session of experts in April 1999, has recommended the constitution of research commissions in order to study and harmonize different linguistic systems. However, they retain the principle to transcribe five main tones (i.e. low, high, rising, downward, and medium tones).

In my view, besides the argument of overloading texts, the transcription of tones in Gabonese languages is difficult in many ways:

- There are various types of tones in languages, and sometimes even the speakers do not pay attention to them.
- In tone languages, most modulated or complex tones (i.e. rising, downward, high-lowed, etc.) are simply the outcome of successive punctual ones (i.e. low, high, and medium). In other words, modulated tones do not really exist on their own,

they are the result of the combination of punctual tones. Therefore, in my view, their transcription in the orthography has to be automatically put aside.

- One of the arguments in favour of the transcription of tones is that they play a distinctive role in the vocabulary. People that reject the transcription of tones estimate that the context or the environment of a word allows the identification of the melodic pitch during the pronunciation. In fact, the distinction of words, which are morphologically identical but semantically different through their tonal structure, is easier in a situation of communication. When a word is put in a context (e.g. in a sentence) its semantic value can be easily identified. Therefore, the context can play a role as relevant as the transcription with regard to tones.
- Even the transcription of tonal markers does not always solve the matter as regard the semantic distinction between words, as one can see in a language such as Civili¹⁰:

cikóóku "relics" vs. **cikóóku** "Punu dance"
múnù "mouth" vs. **múnù** "any form of entrance"

Source of examples: Soami (unpublished)

As one can see, **cikóóku** and **múnù** are identical at the segmental as well as at the supra-segmental level. Nothing in the written form can allow the reader to distinguish one form from another. Only the context of use will be helpful in that regard.

It is necessary to note that the phenomenon of tones does not exist only in Bantu languages. In fact, many Asian (e.g. China, Japan) and American languages have the same phenomenon. In

¹⁰ Civili is a Bantu language spoken in Gabon

some cases, these languages are even more complex than Gabonese languages. Many countries with tone languages have chosen orthographical writing systems without tones. It is the case in Southern Africa (e.g. Namibia, Botswana, South Africa, Swaziland, etc.) and Eastern Africa (e.g. Kenya, Tanzania, Sudan, etc.). In Central Africa, many countries like Rwanda, Burundi, Uganda, DRC, etc. have writing systems that do not include tone markers. However, most of these countries are more advanced than Gabon in the development process of their local languages. Languages of great expansion in Africa such as Swahili, Hausa, Ciluba or Lingala are also written without tone markers. The written tradition in these languages is secular and very well established.

In Africa, some languages have known different situations as far as the transcription of tones in the writing is concerned. A language like Tswana for instance was written without tone markers by missionaries and colonial administrators up to the 20th century. In the middle of this century, tones were introduced in the orthographical system of that language (Cole & Mokaila, 1962). However, despite the use of tonal markers in Tswana literature, that disposition did not resist the negative response of the majority of people. That is why Tswana again became a language without tonal markers since 1970.

The following figures represent written texts in some African languages:

Figure 6: Example of texts in Hausa
Source: Ndinga-Koumba-Binza (unpublished)

The image shows a screenshot of the BBC Hausa website. At the top, there are navigation links for 'BBC HOME PAGE', 'BBC NEWS', and 'BBC SPORT', along with 'Contact Us | Help | Text Only'. The main header features the 'BBC HAUSA' logo. Below this, there are several sections:

- Shafin Farko**: A section for the first story, 'Game da mu Shirye-shiryen Ma'aikata', with a sub-headline 'Labarai cikin sauti'. The text discusses a game between Nigeria and Iraq, mentioning a goal by Shugaba Obasanjo.
- GIDAN REDIYON BBC**: A section for radio programs, listing times like '05:30 - 06:00' and '13:45 - 14:15', and a program 'A Najeriya har yanzu jama'a na ci gaba da tofa albarkacin bakin su game da sabbin Jam'iyyun siyasar da hukumar zaben kasar ta yi wa ragista'.
- OTHER BBC SITES**: A list of regional sites including 'AFRIQUE', 'BRASIL', 'GREAT LAKES', 'PORTUGUES', 'SOOMALIGA', 'KISWAHILI', 'BBC WORLD SERVICE', and 'LEARNING ENGLISH'.
- Weather**: A widget with a 'Select a City' dropdown menu.
- Program Listings**: On the right side, there are boxes for 'THE WORLD'S TOP TEN', 'WASA-YA-YI-WASAL!', 'THE STORY OF AFRICA', 'FOCUS ON AFRICA', and 'NETWORK AFRICA'.
- Breaking News**: A section titled 'Kenya seeks two missile suspects' with a sub-headline 'Kenyan police release photos of two men suspected of carrying out last month's failed missile attack in Mombasa on an Israeli airliner'.
- Also:** A list of other news items like 'Morocco 'al-Qaeda' trial to resume' and 'Ivory Coast mobilises youth'.
- Contact Information**: At the bottom right, there is contact info for BBC Hausa, including an email address 'hausa@bbc.co.uk' and a London address.

NB: The Hausa, Afro-Asian language, has three main tones; low, high and downward. The intonation also has a great place in the language, however none of these prosodic phenomena are graphically represented in the writing.

Figure 7: Example of texts in Kiswahili
Source: Ndinga-Koumba-Binza (unpublished)

The screenshot shows the BBC Kiswahili website interface. At the top, there are links for 'HOME PAGE', 'BBC NEWS', and 'BBC SPORT'. The main header reads 'BBC KISWAHILI'. Below this, there are navigation options like 'Ukurasa wa kwanza', 'Mita bendi', 'Michezo', 'Elimu', 'Historia', 'Anwani', and 'In English'. A 'KWA SAUTI' section lists audio programs at 03:00, 04:00, 15:30, and 17:45. A 'WEATHER' widget shows 'CHAGUA MJI'. The main content area features a large article titled 'Dira Ya Dunia' with the sub-headline 'Habari Mpya 03:00, 04:00, 15:30, 17:45 GMT'. The article text discusses Iraq's weapons and the situation in Burundi. To the right, there are several smaller sections: 'Nominat your favourite song', 'Kenya baada ya MOI', 'Makala ya Somo juu ya Elimu ya Uhusiano wa kijinsia', 'Makala ya Dunia Ya Watoto juu ya Haki za Mtoto Duniani', 'THE STORY OF AFRICA', 'NETWORK AFRICA', and 'FOCUS ON AFRICA'. The date 'HABARI MPYA: 08/12/2002' is displayed at the bottom of the main content area.

NB: Kiswahili was written with the Arabic writing system up to the end of the 19th century. Nowadays it is written with Roman characters, but it does not include any supra-segmental symbols in the orthographical writing.

In my view, any language can be learnt without transcription of supra-segmental markers. In Gabon, many Europeans have learnt several local languages without learning tones. The learning of tones as any prosodic phenomenon is generally not a conscious endeavour for the learner as it comes with practice and the habit of

speaking the language. Therefore, it is possible to learn the language without mastering tones. In addition, if Gabonese languages have to be included in the educational system they will be learned at all levels of the education (i.e. from the pre-primary to the university). In this case, it will be difficult to ask teachers to instruct students about tones when they (i.e. the teachers) do not know with certainty what tones are all about (all the teachers are not linguists).

Finally, the last argument against the transcription of tonal markers in the orthography of Gabonese languages could be the fact that tonal systems of languages are permanently changing. The more a language develops, the more its tonal system changes and gets simplified. Languages lose their tones as they evolve or come in contact with other linguistic systems. If tonal markers are included in the orthography, they have to be regularly readapted every time there is any change in the language.

5.4.4.8. Conjunctivism or disjunctivism

For the writing of African languages, two types of writing systems have been used up to the present; the conjunctive and the disjunctive writing systems. These two traditions are based upon different linguistic assumptions. "Conjunctivism and disjunctivism are two different traditions of word division. The difference between these two systems concerns the status of certain linguistic elements. In conjunctivism, linguistic elements are joined to the following elements, whereas in the disjunctivism they are written separately" (Van Wyk, 1995:83).

The following example is an illustration of a Fang sentence written in both ways:

| Conjunctivism | Disjunctivism |
|---|--|
| Ebyom menga lik nda bisefe ane menga lik byë | E byom me nga lik nda bi se fe an me nga lik byë |
| English translation: <i>The stuff I left at home is no longer as I left it.</i> | |

I agree with Van Wyk (1995:83) that “the choice whether to use the conjunctivism or disjunctivism is purely a matter of orthographical convention. However, one should also be aware that phonological structures of languages are, to a certain extent, influenced by their natural predispositions”. In that regard, Van Wyk gives the example of Nguni languages where vowel deletion and vowel coalescence occur across word boundaries, making difficult the orthographically separation of some words in the concerned languages. Some researchers, like Prinsloo¹¹, consider that disjunctivism is better, particularly for the corpus, as it allows the lexicographer to have a more appropriate frequency and concordance list.

For Van Wyk (1995:84) “conjunctive writing has the benefit of having been championed by Doke, who attempted to justify it on phonological grounds, whereas no authors have provided a theoretical basis for disjunctivism”. As a result, conjunctivism has been accepted uncritically by many African linguists as the only linguistically correct method, on the unjustifiable assumption that there is a one-to-one correspondence between conjunctively written, phonologically defined “words” on the one hand, and grammatical words on the other. According to Van Wyk, “it has been proven conclusively that neither conjunctivism nor

¹¹ Personal communication.

disjunctivism is linguistically justified and that, like disjunctivism, conjunctivism is merely an orthographical convention”.

As far as Fang is concerned, at the actual stage of the development of the language, it is difficult to take a definitive position regarding the orthographical method of writing. In my view, the most important point is that the chosen written method must be the only one used by everyone for the orthographical writing of the language. The tradition of the users should also be taken into account. In fact, some people are already accustomed to write some Fang words in a certain way. They do not care about the way these words are written, they just use them as they learned them. Therefore, whatever the method, these words should be left as they are already used. That will facilitate things for the users. However, in my view, each method should be used carefully. In the conjunctive method for instance, it is difficult to make a clear distinction between words and sentences in a translation dictionary for instance.

The Fang word **madzi**, which corresponds to the English sentence “*I eat*”, does not give the opportunity to the user to know what the correspondence is between *I* and *eat* in Fang. On the other hand, the distinction between **më** “*I*” and **adzi** “*to eat*”, which logically should become **më adzi** to say *I eat*, does not correspond to the way people speak. In that situation, conjunctivism will be the best method, because it corresponds to the way people speak, but at the same time, the user should be aware of the fact that **madzi** is the combination of **më** (*I*) + **adzi** (*to eat*).

The lexicographer may also be confronted with another type of situation in Fang, in which disjunctivism seems to be the best

method. If **bëngakënda** (they went home) can be considered as a Fang word according to the conjunctive method, writing it in the disjunctive method (**bë nga kë nda**) has two advantages:

1. It allows the user to make a clear distinction between different components of the sentence; **bë** (subject *they*), **nga** (mark of the preterit), (**a-**) **kë** (verb *to go*), **nda** (complement *house* or *home*);
2. Keeping the words separately immediately shows the user that he/she is dealing with a sentence and not a word, and this presentation does not interfere with the way people speak. The orthography corresponds to the way the sentence is pronounced whatever the method used.

Personally, I adhere to the disjunctive method and I recommend it for the proposed model because the dictionary should have both knowledge-directed and communication-directed functions. Writing words separately will allow the users to have a clear idea of the item categories (e.g. noun, verb, adjective, subject, article, etc.) composing a sentence, and consequently they can be able to use these items in other contexts. However, if the disjunctive system has to be the general method, the conjunctive system can be preserved for items already traditionally written in that method by users or when that method better shows the way people are speaking, as demonstrated with the item **madzi**.

5.4.5. Concluding remarks

The development process of Gabonese languages remains stuck because of many reasons. The choice and the use of a unique alphabet is one of the first and most important steps that have to

be reached in order to really launch that process and expect better results. Despite all proposals made by researchers in that regard, political decisions have to be made not only in order to organize research units (e.g. lexicographical units), but also to apply the outcomes of such research. The Gabonese government should decide on the adoption and the teaching of a standard alphabet in schools. With regard to the transcription of tonal markers in the orthography, in my view, the example of other countries should inspire Gabon. As it was the case in many countries, people have always systematically rejected systems with tonal markers and Gabon will not be an exception. That is why I suggest that tonal markers should not be included in the orthography of Gabonese languages. They should only be used for phonetic transcriptions, particularly in linguistic works.

With regard to the writing method, I suggest the use of the disjunctive approach for the orthographical writing of Fang. However, that method will just be a more general tradition, because in some cases the conjunctive approach could also be used. The main purpose is to make things easier for the users. Any method or approach used by the lexicographer must always have two main objectives: the development of the language and the empowerment of the user.

5.5. Dictionaries and the standardization of Fang

5.5.1. Introduction

To understand the role dictionaries can play in the standardization of Fang, it is important to be aware of the authority of the dictionary in the society. Indeed, dictionaries have been used for centuries in the acquisition and correct usage of language. Early

dictionaries were compiled in the present-day Iraq 1000 BC. These dictionaries were assisting Assyrians, who came to Babylon, to understand Sumerian signs. Dictionaries consisted of Sumerian signs with their Assyrian translations and had an educational function as instruments in language teaching (Gouws, 2000:4). In Gabon, the first dictionaries had the objective to facilitate the evangelisation of the population and allow colonial administrators to communicate with local people without interpreters. Since the advent of the modern education, the influence and authority of the dictionary became increasingly visible in the Gabonese society.

5.5.2. Role and Authority of the dictionary

5.5.2.1. Introduction

The term "authority" here refers to the power the dictionary can have amongst users by influencing their actions, opinions, or beliefs. During the Middle Ages, when lexicography was still at the beginning of its growth, the use of dictionaries was restricted to small user groups. The access ordinary people had to dictionaries increased their authority. "People argued that if something is printed in a dictionary, it has to be the truth. Dictionaries were regarded as prescriptive sources emphasising the correct language usage" (Gouws, 2000:5).

According to Wells (1994:92-93) "the corollary idea that the dictionary operates as a conservative force on language is prevalent, even among scholars of language". Sterling A. Leonard (as cited by Wells, 1994:93) said: "certainly, the very conservative larger dictionaries may be trusted to delay acceptance long enough. Once an expression is entered as 'colloquial', that is, acceptable for all informal uses by cultivated people, it should be unquestioned in

such use". "The desires to stabilize the language, then, and the notion that the dictionary can be the means to achieve that end, have been linked in our culture since the late seventeenth century" (Wells, 1994:93). Besides the fact that they serve for the learning of new languages (e.g. bilingual or multilingual dictionaries), dictionaries also allow the improvement of knowledge in a language one already speaks (e.g. monolingual dictionaries), they allow users to know better the society and life through named things (e.g. encyclopaedic dictionaries), and to understand certain specific or special domains (e.g. LSP dictionaries).

Dictionaries even allow the projection in sacred or initiatory universes (e.g. dictionaries of rites or secret societies), or to become skilled at universal wisdoms (e.g. dictionaries of proverbs and common sayings). In other words, dictionaries are the projection of the society, the culture, and the life through the language. In Gabon, the authority or influence of the dictionary is particularly visible among (1) the speech communities, (2) the learners, (3) the educational system, and (4) the socio-professionals.

5.5.2.2. Influence of dictionaries in the speech community

By "speech community", I mean a group of people speaking the same language and who use the dictionary particularly to improve their knowledge in that language. By the fact that the dictionary is a representative of the language, it is supposed to take into account the right or good usage of that language, and to put the best options at the user's disposal. With that in mind, the dictionary is considered by the speech community as the ideal source to improve its competence in language at all levels (i.e. vocabulary,

orthography, pronunciation, semantic, conjugation, grammar, etc.). Because dictionaries have been compiled and used for centuries, it gives them a powerful position among all other instruments or books of knowledge. Indeed, the longevity of the dictionary is an important factor in the authority it has acquired among users. That can be illustrated by the fact that dictionaries compiled by early publishing houses (e.g. Oxford and Larousse) are more appreciated by users. The penchant of users for dictionaries from old publishing houses does not mean that these dictionaries are inevitably better than those from other publishers, their age and constancy grant them a certain authority among users.

It is, in my view, the same reason that explains the fact that some religious books as the Bible or Koran have, in general, a bigger authority among users compared with other books of knowledge. In all human societies, the mastering of language has always granted to the one who detains it, a certain authority among the rest of the members of the speech community. Knowing the language and mastering it better than the others has always been the sign of certain erudition. The language being the representative of the culture, its authority is therefore somehow transferred in the dictionary.

In Gabon, heads of villages, clans or tribes were generally those who knew how to speak, who knew the language. Their mastering of the language was a sign of their knowledge and wisdom, even their power, and the community recognized that by the respect and the responsibilities they were given. Thus the Gabonese society recognises a certain authority for one who masters the language. Even if dictionary culture is not yet well-

established in Gabon, there is no doubt that when Gabonese languages will be taught or used as medium of instruction, dictionaries in Gabonese languages will be considered as holders of the language and culture, as it has been the case all over the world. This will give these dictionaries a definite authority.

5.5.2.3. Dictionaries and learners of a new language

By “learners”, I mean people who use a dictionary to learn a new language. In that case, they often use a translation dictionary. The main reasons that give an authority to the dictionary among the speech community remain valid as far as learners of a new language are concerned. However, contrary to the speech community that can be “active users¹²”, learners of a new language are generally “passive users¹³”, because they take the information from the dictionary as it is. They do not have enough competence in the language to do the exegesis of the dictionary. The major part of knowledge these learners will get in the language will be coming from the dictionary, and it consequently becomes the main (sometimes only) source of their linguistic competence. In that situation, the dictionary takes an authoritative place among users.

5.5.2.4. Dictionaries and educational system

According to Dubois & Dubois (1971:11), dictionaries have the right of veto on terms used in the language, in the way that they authorize words, constructions, and meanings by including them in the use of the community, by giving them the status of law. At the

¹²As native language speakers, the users can have a critical view about dictionaries existing in their language, because they know the language.

¹³As foreign language speakers, the users do not always have the necessary knowledge to criticise dictionaries existing in the language they are learning.

same time, they can condemn or reject the use of terms. Dubois and Dubois insist on the fact that dictionaries are playing a pedagogical role, because they provide didactic answers to questions. That is why the acquisition of dictionaries is often advised to students in most educational systems over the world. The increasing presence of dictionaries in educational fields can partially explain the compilation of dictionaries in various domains of education, and the compilation of dictionaries giving an important place to extra-linguistic realities (e.g. encyclopaedic and LSP dictionaries).

Nowadays, most teachers and lecturers recommend their students the use of dictionaries, because they do not always have the time in classrooms or lecture halls to give them all the information they regard as important. They often argue that dictionaries, by the fact that they are the outcome of meticulous work and are regularly updated, give students detailed information that the teacher does not always have. A part of the investigation I conducted, in Libreville (January & February 2001), among librarians, booksellers, students and teachers, shows that the dictionary is the most requested and consulted book in study and research areas (i.e. libraries, bookshops, study centres, etc.). That situation is justified by the fact that dictionaries are nowadays compiled in almost every domain or subject studied in schools and universities.

There exist dictionaries about the ordinary language as well as special fields, like medicine, mathematics, linguistics, etc. All this shows that the authority of the dictionary is nowadays unascertainable in the educational field in Gabon. When dictionaries

will be regularly compiled in Gabonese languages, they will be demanded and frequently used in the educational field. Therefore, their influence in the use of these languages is evident.

5.5.2.5. Dictionaries and socio-professional fields

According to Nyangone Assam (2001:190), dictionaries are highly considered products. The experience shows that the dictionary is the book most frequently found in offices and administration desks. Computing is a good illustration regarding the "imperialism" of the dictionary in the socio-professional field nowadays. Indeed, computing, particularly data processing is nowadays an omnipresent domain in economical and professional activities, because it becomes fundamental for all socio-professional categories that do not want to be "swallowed" or "killed" by globalization, to get indispensable knowledge in computing. Besides the practice in front of a computer, many people want to know more about computers and related activities (e.g. internet). In this regard, dictionaries are very useful for them.

In Gabon, there is a college of education specialized in the training of personal assistants¹⁴. In that college of education, students at the end of their training must be able to use computers or typewriters, speak easily and correctly (even on the phone), write administrative or formal letters, etc. Most of this work is done with the help of the dictionary.

5.5.2.6. Concluding remarks

When I am talking about dictionaries and standardization I particularly mean why and how dictionaries can help in the

¹⁴ Ecole Normale Supérieure (ENS) = High Education School

standardization process of a language like Fang. Therefore, it was necessary to demonstrate the influence and the authority of the dictionary among all categories of potential dictionary users. Even if dictionary culture is not yet strongly established in the habit of Gabonese people, it is on the other hand evident that the first target group (i.e. young Fang people) concerned with the compilation of a French-Fang translation dictionary, is very much aware of the usefulness of the dictionary and they will strongly rely on it for academic purposes. Therefore, the influence and authority dictionaries do have among them is evident, and they will certainly help to fix the standard Fang dialect among this target user.

5.5.3. Dictionaries and the development of Fang

5.5.3.1. Introduction

The role dictionaries can play in the standardization of Fang is not only possible because of their authority among users, but also because of the role they can play in the preservation, and the development of that language. Rey-Debove (1971:13) observes that the dictionary is a book that describes the language by the lexical approach. It includes words in general, including grammatical ones. She adds that the dictionary is also a grammar, through the grammatical data it gives about lexical or grammatical units.

Rey-Debove's observation allows us to see that a dictionary includes many aspects of language, not only through its lexical representation, but also by all other structural elements that compose it (i.e. pronunciation, part of speech, semantics, morphology, etc.). Many researchers (e.g. Lazore, 1993) confirm that dictionaries can participate in the development of languages by admitting that revitalizing any language requires modernizing the

vocabulary, publishing a standard writing system, and developing methods to incorporate new words into the language. The existence of modern dictionaries and grammars makes the work of standardization much easier.

The main objective of this section is to emphasize some of the important relations that can exist between a dictionary and a language, in particular with regard to aspects related to the conservation, the promotion and the development of the language. Particular attention will be given to the relations between the dictionary and contemporary Fang and the writing and reading of Fang. The role the dictionary can play with regard to the use of Fang as academic subject and language of teaching, the acceptance of the standard dialect of Fang, the pronunciation of Fang, and the history of Fang will also be emphasized.

5.5.3.2. Dictionaries and contemporary Fang

By "contemporary Fang", I mean the way this language is or was used (in written and/or oral form) at a specific period. That period can be current, past or future. Languages are, as communities who speak them, dynamic structures that evolve and change time after time, and according to the areas. Dictionaries, as symbols of an advanced culture, are also part of that culture (Dubois & Dubois, 1971:8).

The way Fang is currently spoken is certainly not the same as thousand years ago. This language has recorded mixtures, adaptations, evolvments, etc. according to the needs of the population or according to new situations they met. There is no doubt that an expression such as **misini a mikon**, which means *train*, did not exist in Fang before the contact between the Fang

population and Westerners. Any dictionary compiled in Fang at a specific period represents the Fang of that time. Even when Fang is not spoken anymore it remains alive in the dictionary. Therefore, dictionaries can reflect the diachronic progression of Fang. Dictionaries can also show how Fang has been upgraded as a result of new epochs and events. In fact, with the new technological evolution and communicational means, globalisation and multilateral exchanges, it becomes more and more necessary, even compulsory for Fang to integrate these new realities that relate to the environment of the new world. Realities such as computing, Internet, Email and website, for instance, must necessarily be named in Fang today, if that language wants to become a subject and language of teaching. In this regard, the proposed standard dictionary model will play a decisive role in order to allow Fang to become a modern language, capable to be used in any context or need of the modern world. In fact, that dictionary will give an important place to terms related to modern human activities or notions. It will then help Fang to become a practical language that can be used in any communicational situation. That gives credit to Dubois & Dubois (1971:8) who declare that the dictionary is the book that detains the common treasure of the people, which is the language. The proposed dictionary will be, somehow, the memory, and the favourite reference for linguistic and cultural knowledge in Fang.

5.5.3.3. Dictionaries and the reading and writing of Fang

As already said in a previous section of this work, the dictionary can help the development of Fang. However, that development will only start once some basic steps have been taken. The dictionary has

the purpose to establish a bridge between what the user knows and what he/she does not know (Rey-Debove, 1989:2817). One of the most important changes Fang needs is to move from an oral language to a written and read one. The proposed standard dictionary model can help it to pass from one stage to another.

In previous sections, I expressed the necessity to choose and use one unique writing system for the orthographical transcription of Fang. That step is a *sine qua non* for the writing and the reading of the Fang before one can think of any further development. Besides the choice of the standard writing system, the proposed standard dictionary model will have the role to promote and popularise that writing, in order to give users the habit to read and use Fang through the dictionary. The proposed standard dictionary model will not only promote the standard dialect, but also the orthography or spelling. If words are written in a certain way in the proposed Fang standard dictionary model, users in their everyday use of the language will write them in the same way. When a word is in the dictionary, it is written and read by the users as presented in the dictionary. As the proposed standard dictionary model will be present in academic fields, it will be an important educational tool. Therefore, it will help to speed the validation, the acceptance and the use of Fang orthography.

5.5.3.4. Dictionaries and Fang as an academic subject

The proposed standard dictionary is meant to fulfil the needs of young Fang people who want to learn the language of their parent. Its compilation will be the answer to the need of the target group to communicate fluently in Fang and it will have the purpose to satisfy the needs of that specific group of users. Through Fang, they want

to keep alive their culture and develop their identity. Therefore, the proposed standard dictionary model can be an instrument in the teaching of Fang. As any modern dictionary, the proposed dictionary model will include and present all aspects necessary for the learning and improvement of Fang:

- Lexically, terms and expressions included in the proposed dictionary model will be retrieved from a corpus compiled with linguistic data of Fang. These terms and expressions will be selected according to their frequency and importance for the Fang people, or according to the linguistic needs of young Fang people.
- Semantically, the proposed dictionary model will present all pertinent relations (e.g. homonymy, polysemy, synonymy, etc.) that will allow the user to seize the meaning and context of use of the included lexical units. Consequently, the proposed dictionary model will try to give the user all types of data that can allow him/her to learn and utilize Fang correctly.
- As far as the Fang varieties are concerned, the proposed dictionary model will take into account all particularities of the language (i.e. differences and similarities), even if an important place will primarily be given to the standard variety (i.e. Fang-Ntumu). Thanks to the proposed dictionary model, the user will have an idea about Fang dialects.
- Morphologically, paradigms, derivations, components, and combinations of words in the language are also parts of the lexicographical treatment of some dictionaries. Diachronic dictionaries, for instance, can give the origin of words through

the morphemes that constitute them. Even if the proposed dictionary model will not be a diachronic dictionary as such, it will present the origin of some words where their original denotation will be necessary to understand their current meaning in Fang.

The compilation of the proposed dictionary model will allow the user to have most aspects related to the functioning and use of the Fang. That will facilitate the learning of that language by students in schools. The *French/Fang standard translation dictionary* will also be a didactic instrument from which Fang teachers could retrieve the basic tools for teaching the language.

5.5.3.5. Dictionaries and Fang as language of education

At the current stage of its development, one can ask if Fang can be a language of education at all levels. The matter here is not whether Fang has or does not have the terminology to be used as language of education, but the real question is how to create the conditions that will render the teaching in Fang a reality. There is no doubt about the fact that if modern realities have been named in other languages, they can be named in Fang too.

Let us not forget the fact that even in Western societies, many current realities were unknown before, and consequently not named in their languages. They had to update their language according to new realities in order to name them. Therefore, there is no reason for concern regarding the lack of Fang terminology to name modern and academic realities. However, it will be naïve to ignore the fact that the designation of several modern realities is not regular or evident for many Fang speakers. Even when terms

exist in the Fang, many people do not know them. It is the duty of the compiler of the French/Fang standard translation dictionary to try to satisfy the needs of these speakers. Indeed, to teach and/or learn mathematics, computing, physics, etc. in Fang, teachers and students will need to learn or familiarize themselves with Fang terms used in these domains and all academic subjects. The compiler of the proposed dictionary model has the responsibility to make that data available to them, not only to help in teaching or learning the language, but also in order to help the Fang to develop and update itself with new realities by the creation and/or loan of terms designating these new phenomena. In this regard, the proposed dictionary model will be an important instrument in the process of using Fang as language of education.

5.5.3.6. Dictionaries and the development of the standard form of Fang

The standardization of Fang is the process by which the standard form of that language is established. The standard form in question is one of the varieties of the specific language, which will be chosen and accepted as norm by the speech community. It is impossible for Fang-Ntumu to become the standard form of Fang, even when it does match with all the criteria, if the speech community does not accept and use it as reference. It is at this point that the proposed dictionary model will play a very important role. The structural properties of the standard language are the areas where the influence of the dictionary is most visible (Gallardo, 1980: 62).

The proposed dictionary model will influence the acceptance of the Fang-Ntumu as standard dialect by the speech community in many ways. Zgusta (1980) has stressed four types of dictionaries

that can play a role in the genesis and the development of the standard form. By using features highlighted by Zgusta, the proposed dictionary model plays an important role in the genesis and the development of the standard in Fang according to the following ways:

- The proposed dictionary model will aim at creating a written standard. It will have standard-creating features, as it will be the outcome of cultural activities in the speech community. For instance, the need to teach and introduce new concepts in Fang will show the necessity to write them in the language. In that situation, the proposed dictionary model will introduce into the standard form of Fang terms lacking in it but present in the source language, i.e. French. The compilation of the proposed dictionary model supposes that there are or soon will be Fang speakers that have received various types of education in French and who will need the help of the dictionary in order to write Fang. On the other hand, the fact that the proposed dictionary model will present in Fang an important lexicographical treatment of equivalent items, it can help to achieve the purpose of fostering Fang as a written standard language. The proposed dictionary model can also be instrument in the survival of Fang by allowing it to introduce itself as a written language. By including many important aspects of the other Fang varieties, the proposed dictionary model can help those varieties to develop a standard of their own through the compilation of small monolingual dictionaries. As the proposed dictionary model will include the standard Fang and its regional varieties, it will

- allow the users to see the difference between the standard Fang that has to be used in official and academic activities and the other varieties that can be used in any other contexts.
- Another way the proposed dictionary model can develop the standard form of Fang is the inclusion of modern words and expressions. In fact, as in any language, constant change of the lexicon is observable in Fang, because of the incessant influx of new words and expressions. Scientific and other technological fields for instance can be regarded as generator of new terminology. The purpose of the proposed dictionary model is to define terms, which are consistently used in order to communicate effectively. Therefore, it will include the necessary vocabulary that can take Fang to the level of modern language.
 - The third way the proposed dictionary can develop the standard Fang is the inclusion of antiquating or archaising terms and expressions. This dictionary can help to use classical expressions in place of common language. It will try to reverse changes in the language by re-introducing ancient and correct words. For instance, most young people are using borrowed words when they are speaking, because many Fang terms and expressions are not used anymore, even when they exist. That situation can give the impression that the language does not have the appropriate words to name things, and has to borrow them from other languages. The proposed dictionary model will help to bring back these words and re-introduce them in Fang.

-
- The fourth and last way the proposed dictionary model can contribute to the development of the standard Fang is to help the standard dialect to function as a general comprehensible means of communication beyond the boundaries and limits of the other Fang varieties. The standard dialect of Fang will need some stability to remain generally comprehensive for some period. Many users of the proposed standard dictionary will be interested in getting advice about how to correctly use Fang, particularly in educational fields. Such advice will be either implicit by excluding non-recommended items and usages from the proposed dictionary or explicit by expressing labels or usage notes. The proposed dictionary model should also give an important place to the contemporary language. This means that besides the standard-descriptive layer, that dictionary should contain data about obsolete words not used anymore but met in literary works that are still read (e.g. dialectal expressions, colloquialisms, demotic speech, taboo words, etc.).

Zgusta (1980:75) has also stressed means and methods by which the dictionaries can influence the user. These means and methods can be effectively applied to the proposed dictionary in the Fang situation as follows:

- Among several synonyms or other variants, only one or two of them will be listed in the proposed dictionary. That means that the user who will try to communicate in the new standard Fang would use the listed rather than the non-listed variants. The selection of data (lexical units, variants of usage, etc.) is an influential factor because users do check the dictionary

particularly when drafting written texts, and frequently accept what the dictionary says.

- Labels and symbols that give information about the status of expressions (e.g. colloquial, informal, dialect, obsolete, etc.) or that give the indications of syntactic patterns, etc. will also influence the way the user will utilize Fang. For instance if an expression is labelled as informal in the proposed dictionary, the user will probably try not to use it in formal situations (e.g. official speeches, writing of thesis, formal letters, etc.).
- The proposed dictionary will contain short commentaries of various types within the article. These commentaries will influence the value of some words to the user. For instance, the proposed dictionary will have historical slant. Therefore, it will give a historically founded argument for or against a usage. The proposed dictionary will also give encyclopaedic arguments to motivate the selection of a term instead of a competing synonym.

I strongly agree with Zgusta (1980:76) who said that there is yet no empirical data as how far usage notes do influence real usage. However, one can maintain that dictionaries exercise a certain influence on their users, so that they tend to stabilize the norm. Even if the degree of that influence depends on the type of the dictionary, it particularly depends on the willingness of the public to cooperate.

On the other hand, there are areas of lexicography where the influence on the user is stronger because of the overwhelming superiority of the knowledge contained in the dictionary.

Terminological and similar activities frequently performed under the governmental, educational system or similar auspices and the nearly obligatory character of their results are particularly influential in dictionaries. Therefore, there is no doubt that dictionaries do influence the linguistic behavior of their users. They cannot stop change, or cause changes disapproved or not accepted by the user. They can and do stabilize the usage, particularly in the written language and in related formal styles of the spoken language; they do clarify meanings and make them more systematic.

The proposed *French/Fang standard translation dictionary* model will give data about domains of usage and discriminate with regard to the stylistic status of words as well as other types of lexical variation, such as social and geographical, and even temporal words of Fang. All this means that the proposed dictionary model will be a crucial element in the way the frame of reference function of the standard Fang manifests itself, and therefore intimately related to the awareness of the norm attitude among members of the speech community (cf. Gallardo, 1980:62). Because of the arbitrariness of the Fang orthography and the fact that this orthography is not well known by the speech community, it is convenient for the user of the language to have a reference book with unchallenged authority as guide in writing (and later on, even speaking).

As many dictionaries before, the proposed dictionary model will also develop the standard Fang in the way that it will allow the speech community to find and assert its own identity towards technological language models. American English is a good instance in this regard, e.g. through Noah Webster's dictionaries. In fact,

Webster saw that American dictionaries could be the most successful promoters of this way of approaching the standardization process, and not only at the level of the organization of the structural properties of the lexicon. To ensure the rootedness of it, he incorporated all types of American-centered cultural elements in his dictionaries; and, by rejecting British-based literary models that would have forced him to find word origins in British sources, he developed a sort of long-range strictly lexical etymology, ultimately Indo-European. In summary, he created new bases for the unifying and separatist functions to manifest themselves, allowing American to be a valid version of English, different from British (cf. Gallardo, 1980:65). Today Webster' dictionaries are still considered as outstanding symbols of the standard American English.

5.6. Standardization and electronic dictionaries

5.6.1. Introduction

The evolution of dictionaries in the future will probably be connected to that of computers. In fact, the computer, which has changed so many features of the human society, increasingly becomes directly evident in the field of lexicography. Different types of computer software are used in dictionary production particularly for word-processing, database maintenance, textual and grammatical analysis.

With regard to dictionaries, computer systems offer natural language interfaces used in different environments and formats, and generally called electronic dictionaries. With globalization, multilateral exchanges, and perpetual progress of the technology, etc. electronic dictionaries will become accessible and more used by the majority of dictionary users. For that reason, I believe they can

play an important role in the standardization and growth of developing languages as Fang.

5.6.2. Types of electronic dictionaries

There are mainly two types of electronic dictionaries: dictionaries for people and dictionaries for programs (Byrd, 1995:207). Dictionaries for people usually present an interface at a computer terminal or workstation with which the user can request information associated with words. The dictionary may be a stand-alone application, such as an on-line, or a part of a larger system, such as the spelling checker or the terminology dictionary of a translator aid system. All these electronic dictionaries are mainly in three types of formats: CD Rom, On-line and Handheld. Dictionaries for programs, on the other hand, provide information that may only indirectly be made available at the computer system's interface.

Users of systems supporting natural language query, text analyzing, speech synthesis and recognition, or information retrieval never see word information directly at the interface. Other programs rather use that information to implement their applications, and the human users only see the application results. Dictionaries for programs are often built for specialists (e.g. lexicographical units) that are interested in the improvement of their dictionary systems. Therefore, in this section my dissertation will be focusing on dictionaries for people.

5.6.3. Advantages of electronic dictionaries

Contrary to paper dictionaries, electronic dictionaries offer users various sophisticated ways to access the data available in the dictionary. Where the knowledge of the alphabet is needed for the

efficient use of a paper dictionary, the user can access the wanted information in a CDRom dictionary by simply typing the two first letters related to that information. Apart from the fact that the capacity of an electronic dictionary can largely go over 600 Mega octets (Mo), i.e. the capacity of many paper dictionaries, electronic dictionaries are easy to carry because of their weight (they are light) and their format (they are small). Most people are travelling nowadays; they will probably prefer to take an electronic dictionary rather than a paper one.

As far as the target group of the dictionary model I propose is concerned, one knows that most young people are attracted by games. The compilation of electronic dictionaries in Fang could integrate games that will give users the possibility to combine business with pleasure, by learning and playing at the same time, as that is the case in many electronic dictionaries compiled by well-known publishing houses (e.g. Oxford, Larousse, etc.). These dictionaries will encounter a great success amongst young people who will be interested in the acquisition of dictionaries in Fang, particularly when that language will be introduced in the educational system.

5.6.4. Internet and the development of Fang

5.6.4.1. Introduction¹⁵

Mostly known as Internet, the World Wide Web is an open interconnection of computer networks that enables the computers and the programs they run to communicate directly. The great scale and universality of the public Internet results in its use to connect

¹⁵ The biggest part of the information in this section is coming from Microsoft Encarta Encyclopedia 2001.

many other kinds of computer networks and services, including online information and shopping services, via systems called gateways. As a result of all these features, Internet is an ideal means of building a very robust universal information infrastructure throughout the world. The rapid growth of online shops, information services, and electronic business applications is testament to the inherent flexibility of the Net.

The Internet technology was created by Vinton Cerf in early 1973 as part of a project headed by Robert Kahn and conducted by the Advanced Research Projects Agency, part of the United States Department of Defense. Thereafter, Cerf led many efforts to build, scale, and standardize the Internet. In 1984 the technology and the network were turned over to the private sector and to government scientific agencies for further development. By early 2000, Internet access was available in over 200 countries and encompassed around 100 million users. The Internet and its technology continue to have a profound effect in promoting the sharing of information, making possible rapid transactions among businesses, and supporting global collaboration among individuals and organizations.

5.6.4.2. Role of Internet in the standardization of Fang

The lexicographical domain did not stay out of the advantage of the Internet technology. In fact, many on-line dictionaries and lexicographical reference works (e.g. automatic translations) are nowadays available on the Net. Any person having a computer connected to Internet can access these on-line lexicographical works.

In my view, the publication of Fang dictionaries on Internet could be an excellent means to standardize and develop that language. In the future, the use of Internet for personal and academic research will probably increase even in countries where the Internet access is still difficult today. The availability of Fang dictionaries on-line will probably influence the use of that language and enlarge its accessibility to more users. The more a language is accessible to people the more it increases its chances to be learnt by more people. At the same time, the presence of Fang on the Net will oblige the lexicographers working on that language to extend research in Fang.

The construction of websites entirely in Fang, with interactive features or pages, could also play a significant role in the standardization of Fang. That will force people to read in Fang and accustom themselves with the language. People could also exchange in Fang via Electronic mails (E-mails), live or delayed discussions through Internet Chat rooms or forums. These exchanges will accustom people with the writing and the reading of the language.

5.7. Concluding remarks

The dictionary is nowadays and more than before one of the most used instruments for the acquisition and improvement of languages. Beyond its purely pedagogical and educational purpose, the dictionary plays an essential role in the development of languages. Fang, as most of the developed languages in the past, can find through the proposed dictionary model an ideal instrument that will help it to cross the frontier of oral and non-standardized language to the written and standardized one. To reach this objective, the

utilization of any method or means such as a unique writing system or any technological instrument must be entirely part of the process.

CHAPTER 6:

Concluding remarks

The boundaries between the professional activities and the academic field of lexicography are fluid, and several associations, e.g. DSNA (the Dictionary Society of North America, established in 1975), the Lexicographical Society of India (established in 1975), EURALEX (the European Association for Lexicography, established in 1983), AUSTRALEX (the Austral-Asian Association for Lexicography, established in 1990), the Lexicographical Society of China (established in 1992), AFRILEX (the African Association for Lexicography, established in 1997), have been formed to act as bridges between dictionary makers and academic lexicographers. Professional training, regular national and international conferences, seminars and workshops, and academic publications have served to mature lexicography into an independent field with its own principles and practices, purposefully making use of, and qualifying, the findings of other disciplines (Hartmann & James, 1998).

6.1. Introduction

“As a scientific discipline, lexicography has two major components, i.e. a practical component, the practice of compiling dictionaries, and a theoretical component, metalexicography, which presents the theoretical basis for the lexicographic practice. Lexicographic practice is much older than metalexicography. The lexicographic practice developed and became established in a pretheoretical era and environment. However, the last two decades have witnessed the rapid growth of metalexicography, and the endeavour to formulate a coherent general theory of lexicography stimulated the development of metalexicography into a fully-fledged discipline” (Gouws, 2001a:59). The main purpose of metalexicography is to establish a theoretical basis in order to improve the compilation of dictionaries. That purpose is motivated by the desire to be as close as possible to the needs of dictionary users. Thus, this project is a

theoretical model that can be used and applied for the compilation of translation dictionaries in Fang, or any language in a similar situation. However, apart from this purely lexicographic purpose, the present project also aims to demonstrate the role that a dictionary can play as an instrument in the standardization process of a language with several varieties such as Fang. This dissertation as a whole encompasses five main chapters:

- Chapter 1, *Introduction and problem statement*, identifies the motivation for the choice of the present topic, as well as the frameworks (i.e. theories, models, etc.) that I use to build my metalexicographical criteria.
- Chapter 2, *Fang and its dialects*, gives a better knowledge about the concerned language and its dialects. In fact, apart from the historical overview about Fang people, a wide inventory of works that have been implemented in Fang, as well as the choice and motivation of Fang-Ntumu as the standard dialect are presented in that chapter.
- Chapter 3, *Dictionaries and corpora*, highlights the strong relationship between the compilation of modern dictionaries and electronic corpora. More precisely this chapter intends to demonstrate the importance of electronic corpora in the compilation process of modern dictionaries. In this regard, the importance of some software and new methods are also demonstrated.
- Chapter 4, *Aspects of the dictionary structure*, mainly focuses on different structures of the dictionary and how data should be spread and organized in each dictionary component. A

detailed account of the structure of each dictionary component is reviewed.

- Chapter 5, *Dictionaries and standardization*, focuses on the main purpose of the dissertation and demonstrates why and how the dictionary can be an instrument in the standardization process of languages, with specific reference to Fang.

In the following sections, the major aspects of each chapter will be highlighted and new possible developments will be foreseen in the field of lexicography, and particularly with regard to theories and models that have served as framework in this dissertation.

6.2. Introduction and Problem statement

6.2.1. Review

The *introduction and problem statement* deals with the explanation of the topic, the typology of dictionaries and the functions of the proposed dictionary model. The outcomes of that section are that the term *translation dictionary* in the work primarily refers to bilingual dictionaries, and the proposed model is set to provide interlingual translation equivalents between French and Fang. The term *standard* means that the proposed dictionary model should present and treat the language in such a way that it can serve as a basis, example, or principle to which Fang speakers (should) conform, or by which they can be judged in their language use. This dictionary should primarily endeavour to reflect the standard variety of the target language (Fang). It should be normative in its approach. The first chapter also presents the situation of Gabonese languages through different inventories carried out by Guthrie

(1948, 1953 & 1971), Raponda Walter (1960), Jacquot (1978) and Kwenzi Mikala (1987 & 1997). The inventories of Kwenzi Mikala, which are the most recent and include more speech forms (i.e. 62 speech forms grouped in 10 language units), constitute the basis of the classification of Gabonese languages presented in this chapter. In this chapter, the language pair (i.e. French as source language and Fang as target one) is identified and the choice of their status is justified.

This chapter also deals with the determination of the target users (i.e. young Fang speakers), the motivation of the typological model (i.e. monoscopal and hybrid), and the identification of the dictionary functions (e.g. standardization, translation, production, for active and passive user, etc.) of the proposed model. Finally, the first chapter poses the basis of theories, models and frameworks used for the progression of the dissertation. In fact, the general theory proposed by Hausmann and Wiegand (1989) is chosen to identify and build up the component parts and organization of the microstructure of the proposed model, while the approach of Gouws (2000 & 2001) is considered as the right theoretical basis that should guide the organization of outer texts in the dictionary model. As far as the relationship between dictionaries and the standardization of languages is concerned, theories stressed by Zgusta (1971) & Gallardo (1980) are found more adequate to be used as guiding ways in this regard. However, theories and methods used by other researchers such as Bo Svensén (1993), Prinsloo (2000), Rey-Debove (1971), Dubois and Dubois (1971), Béjoint (2000), Kennedy (1999), etc. are also suggested as important references in the first chapter.

6.2.2. Perspectives

The elaboration of metalexicographical criteria for the compilation of modern dictionaries has shown the importance of well-established theories and methods in the field of lexicography. In fact, when planning any dictionary, the lexicographer has to be aware of the latest evolutions in his/her domain. That awareness will allow him/her not only to compile good quality products (i.e. dictionaries), but also to be a modern lexicographer. In my view, the modernity of a lexicographer does not only concern his/her ability to stick with new developments, but also by his/her capacity to update the existing theories and methods according to his/her specific needs. That is why the first chapter of this dissertation tries to define the scope of the topic by giving to each main concept used a perception or vision efficient for the proposed model. However, many of these concepts could have different values in other publications or their current conception could change in the future. In some cases, such changes are already visible through the rising of a broad range of new definitions of certain concepts. In some recent publications, e.g. the concept of translation dictionary is distinguished from that of bilingual dictionary. In fact, some authors (cf. Tarp, 2002) distinguish between translation and bilingual dictionaries. When a bilingual dictionary is defined as a dictionary that has two object languages and provides equivalents in the target language for each word and expression in the source language, a translation dictionary will be regarded as a dictionary designed to assist the user in solving problems related to the translation process, and it can be monolingual, bilingual or plurilingual. The real point is to prove that if bilingual dictionaries are essentially translation dictionaries, the

concept of translation dictionary does not necessarily apply to bilingual ones. This distinction between bilingual and translation dictionaries will probably be taken into account more often in the future by lexicographers in order to avoid any confusion for the laypersons.

Other topics where there could be changes in the future of lexicography are those of dictionary functions and dictionary typologies. These concepts are related because the typology of a proposed dictionary is strongly related to its function(s). In fact, where most dictionaries of the recent past were monofunctional, new dictionaries tend to be more and more polyfunctional. Since most modern lexicographers want to satisfy the needs of their potential dictionary users, their task becomes more and more difficult because of the multiplicity of needs amongst the target users. Despite the fact that they cannot satisfy all their user needs in one dictionary, modern lexicographers are always trying to fulfil most of these needs. That is why most current dictionaries are intended to satisfy more than one type of user needs, and that is consequently leading lexicographers toward the compilation of dictionaries with many features, in other words toward the compilation of hybrid dictionaries. The dictionary model proposed in this dissertation is a good illustration in this regard. It will have at least three main functions: (1) it is planned to satisfy the needs of young Fang people in terms of translation or transfer of knowledge from French to Fang thanks to the presentation of French lemmata with their equivalent items in Fang, (2) it also has the purpose to help the target user group in the acquisition of basic cultural knowledge in Fang through the inclusion of many entries giving

cultural data, and (3) the monolingual dimension of the proposed dictionary and the use of one Fang dialect in the dictionary have the objective to help with the standardization of Fang. That could raise the following question: *are hybrid dictionaries not the future of lexicography?*

Certainly one of the duties of metalexicographers in the future will be to try to give the right answer to this question, and in particular to set the best ways for the compilation of hybrid dictionaries. For instance, the hybridity of future dictionaries does not mean that any feature can be included in any dictionary without taking into account the target group and the functions of the intended dictionary. For that reason, the role of metalexicography will be to stress new theoretical models for the compilation of different types of hybrid dictionaries, in order to identify which type of features can be compatible for each specific type of users, and what could the functions of these dictionaries be.

6.3. Fang and its dialects

6.3.1. Review

The second chapter, *Fang and its dialects*, essentially gives some general information concerning Fang (the people and the language), by focusing on the location of that language in Africa and in Gabon, and through the historical overview of the migration of the people speaking that language nowadays in Gabon. This chapter presents the classification of Fang among other Bantu languages. Fang is identified as member of the so-called Yaoundé-Fang group or the A70 group and is designated by the acronym A75, according to the classification of Bantu languages made by Guthrie. An important overview of the migration of the Fang people in Africa is also

presented in order to have a better idea of the historical expansion of this ethno-linguistic group from the earliest known stage of their origin to their current location in Gabon. This overview shows that the migration of the Fang population is dated around the 18th and 19th century, and the source regions of Dja, Ivindo and Ntem are considered as the last step before the dispersion of the Fang population throughout Gabon. Nowadays, people speaking that language are localised in 4 provinces and more precisely in 17 regions of the country.

The second chapter also presents a clear distinction between language and dialects. This analysis indicates the importance of language as *the most basic means of communication between human beings*. According to some researchers (cf. Anthony Lodge, 1993) language variation derives not only from the different uses to which a language is put, but also from its users, i.e. from variables among speakers, e.g. their geographical origin, social class, ethnic group, age, sex and so on.

This chapter could not be complete without giving a more detailed account of the six Fang varieties in Gabon. That section of the second chapter particularly identifies different areas in which each Fang dialect is localized in Gabon, and most importantly inventories different works (particularly linguistic and lexicographical ones) known in Fang, in general, and particularly in each dialect. From that inventory, one remarks that missionaries and colonial administrators who were not lexicographers as such have authored most of these works. In addition, this inventory shows that Fang-Mekè, Fang-Atsi and particularly Fang-Ntumu are the dialects in which most works have been compiled in Fang. This

observation is an important criterion that has guided the choice or the proposal of Fang-Ntumu as standard dialect for the standardization of Fang. However, the choice of the standard dialect is made according to a wide range of criteria such as: the high degree of declared understanding of the dialect, the numerical importance of the speakers of this dialect, the advantageous geographical position of the dialect, the location of the dialect at the center of economic activities, the prestige of the dialect, the socio-economic importance of the dialect, the existence of written documents in the dialect, the historical expansion of the dialect, etc. Fang-Ntumu is the dialect that meets most of these criteria.

6.3.2. Perspectives

One of the most difficult steps in the standardization process of a language with many varieties is the choice of one of the varieties as standard one. Indeed, choosing one dialect over the others can have many socio-cultural and political implications. As everybody wants their own dialect to become the standard one, people are unwilling to officially adopt and speak a dialect that is not their own. Many people think that speaking another dialect means recognizing a kind of supremacy of the people from that dialect over them. They often think, rightly or wrongly, that their dialect is the best or the right one that deserves to become the standard one. The Fang situation is a very good instance in this regard. Fang-Atsi people regard their dialect as superior and as the best candidate for the standard variety because of the numerous works compiled by missionaries in their dialect. On the other hand, the Fang-Mvaï people are convinced that their dialect is the purest one, because they take for granted that it is less mixed (i.e. has few borrowed

words) than the others. Therefore, to them, it should be the standard dialect of Fang. Fang-Ntumu people use the argument of their number and advantageous geographic and political position (since their dialect is spoken in two major areas where many activities of Fang people are taking place in the Woleu-Ntem province) to plea in favour of Fang-Ntumu as the standard dialect. Fang-Mekè people also use the same argument than the one used by Fang-Atsi people (i.e. numerous religious works compiled in their dialect), but to that they add the fact that their dialect is spoken in the capital city of the country and it is the only Fang dialect in which the translation of the Bible has been done. All these positions show how difficult it is to choose a standard dialect for underdeveloped languages and that urges the necessity for researchers to formulate more scientific criteria for the choice of standard dialects. I would personally like to see the dialect that has less borrowed words to be chosen as the standard dialect. This means that the pureness of the dialect will be a major criterion as far as the choice of the standard dialect is concerned. However, that criterion can only be retained if linguistic studies are undertaken in each dialect. The formulation of new scientific criteria does not necessarily mean that everyone will automatically accept the chosen dialect, but they can constitute powerful arguments that can convince people to accept and adopt the chosen standard dialect, particularly if it is different from their own one.

Another reason that explains why people reject the other dialects is the fear to see their own dialect disappear, as their dialect represents their identity, and constitutes their speech community. They would be more willing to adopt another dialect if

they could be sure that their own one will not disappear. In other words, they will be able to accept and speak the chosen standard dialect (if it is not their own) if they know that the choice of a standard dialect does not necessarily mean the disappearance of other dialects and that they will still be able to speak their own dialect in other circumstances than in official functions. Lexicographers can play an important role in that process to reassure the dialect community by compiling not only standard dictionaries, but also dictionaries with non-standard varieties, or by including in standard dictionaries items from the other varieties. In such a case, the lexicographer must make sure that the user can distinguish the standard items that have to be used in official and academic functions from those from other dialects that he/she can use in any other context. The compilation of dictionaries in non-standard dialects or the inclusion of non-standard items in dictionaries can be an important argument to convince people that using the standard dialect does not mean the disappearance of their dialect or the superiority of the standard dialect over the others.

In the Fang situation, the choice of Fang-Ntumu is not necessarily the perfect one, as the other dialects also have their own advantages. Therefore, that choice could be contested, as it remains subjective despite all the identified criteria. However, according to my point of view the number of speakers, the advantageous geographical situation and particularly the important role of Fang-Ntumu in socio-economical and cultural activities in the Woleu-Ntem province are important strong points for the choice of that dialect as the standard one. To make that choice effective and acceptable to people from the other dialects, many socio-cultural

issues should be taken into account. One of them, maybe the most important, is the social behaviour of Fang-Ntumu people. In fact, people from the other dialects often reject Fang-Ntumu because they think that people from that dialect regard themselves as superior with a condescending attitude toward people from the other dialects. This aspect may be non-scientific and subjective, but it cannot be ignored since the choice of a standard dialect involves many extra-scientific and subjective aspects. In my view, the attitude of Fang-Ntumu toward the other members of the Fang community in the future could be determining for the acceptance or the rejection of that dialect as the standard one. If their attitude is humble and friendly, Fang-Ntumu will more readily be accepted as standard dialect than if they remain conceited, as many people seem to think.

As Fang, the large majority of Gabonese languages are in the situation where they need to be standardized and thus really launch their development process. Therefore, standard dialects have to be chosen in each language. This means that the difficulties highlighted in the Fang situation will also apply to the other Gabonese languages. Concerning the attitude of people from the chosen standard dialects, a specific didactic program could be included in the educational system (e.g. in civic instruction) that will educate students to show the right attitude toward the other members of the community (particularly the attitude of people from the standard dialects), and the necessity for the other members of the speech community to accept the chosen dialect for the development of the language. As one can imagine that the participation of the Government will be important to reach that objective, since it is the

only authority that has the ability to include these aspects in the educational system.

The development of Fang also means the development of research in different aspects of the language, particularly research that can allow a better knowledge of that language. In this regard, linguistic studies should be encouraged in all dialects. These studies will allow researchers to know with certainty how the language is working and what aspects of the language could be useful for the language learners and how they could be efficiently included in Fang dictionaries. For instance, linguistic research can study the rules of agreement in Fang. Even if this research is done in more technical ways, lexicographers can use the result as additional data and include them in dictionaries in the way that can be understood by anyone. Linguistic research could also allow the lexicographer to have a broad idea of differences between dialects. For instance, phonological, morphological and semantic studies could lead to knowledge whether certain categories (e.g. nominal classes, parts of speech, stems, etc.) are identical in all Fang dialects. All these examples show the necessity for a better collaboration between linguists and lexicographers.

6.4. Dictionaries and corpora

6.4.1. Review

Chapter 3, *Dictionaries and corpora*, focuses on a practical aspect in the compilation process of dictionaries. In fact, nowadays most dictionaries are compiled with the assistance of electronic corpora that are useful for the lexicographer not only for the identification of lemma candidates, but also for the lexicographical treatment in the microstructure. In this chapter, the concept of corpus is described

according to different fields. When the corpus is generally described as *a very large collection or a body of words, usually stored in computer format*, in the language sciences it is defined as *a body of written text or transcribed speech, which can serve as basis for linguistic analysis and description*. Some authors (cf. Landau, 2003) consider that when one speaks of a corpus in lexicography, it mainly refers to an electronic corpus, often one containing a vast number of words from many different sources.

Chapter 3 also tries to identify different stages in the process of corpus building, from the selection of sources through to mark-up and assigning annotation to assembled texts. Particular attention is paid to computational methods and tools designed to assist the various lexicographical tasks, including the preparation of lexicographical evidence from many sources, the recording in the corpus of the relevant linguistic information. Mention is also made of the population group from which the data (oral & written) is retrieved, with the objective to compile a standard French/Fang bilingual dictionary. An important issue concerns the role of data processing (computers, software and networks) in the compilation of corpora, as confirmed by many researchers (cf. Prinsloo, 2000), who consider that compiling and querying electronic corpora has become a *sine qua non* as an empirical basis for contemporary linguistic research, because good dictionaries increasingly base the compilation of both their macro- and microstructure on electronic corpora. As the proposed dictionary model is intended for an African language (i.e. Fang), chapter 3 is the venue in which some approaches and experiences in corpus building in African languages are described. That is why for each step described in the

compilation process of electronic corpora in African languages, chapter 3 proposes how these steps could be applied to the Fang context.

Chapter 3 also gives a more general account of electronic corpora. In fact, apart from the presentation of different types of corpora, different important phases of the compilation process are presented in a more detailed perspective, particularly aspects related to the material acquisition phase and the material preparation phase. The general acquisition focuses on the criteria that guide the material selection; these criteria are essentially based on the authenticity of the sources, the areas covered by the sources, the suitability or usefulness of the sources, the representation of all language registers, and finally the timing. The general acquisition phase also identifies different sources (i.e. primary, secondary, and tertiary) in which data can be collected for a French/Fang standard translation dictionary. As far as the material preparation phase is concerned, it focuses on how the collected and selected written and spoken material have to be transferred from the original medium to computer-readable form. In this regard, apart from the usual keyboarding method, other types of data transfer methods such as downloading (from Internet) and scanning of printed matter are also discussed. The material preparation also gives an important place to the text encoding phase which consists of any combination of word tokenisation, part-of-speech tagging, lemmatization, syntactic parsing and marking up. A necessary regular updating phase of the corpus, with some explanations about how that updating should be done, is also recommended in this chapter.

Finally, chapter 3 presents a discussion of the relation between corpora and standard dictionaries. In fact, it has been highlighted that thanks to electronic corpora, lexicographers can easily distinguish items that belong to the standard dialect from those that belong to the other dialects. As a result, standard dictionaries can be produced and revised much quicker than before, and can thus provide up-to-date information about the standard language. In the same vein, definitions can better reflect the standard language since a larger number of natural examples from the standard dialect can be examined. Instances extracted from the standard corpus can easily be organized into more meaningful groups for analysis. Corpora play a great role in the compilation of standard dictionaries as they enable lexicographers to rapidly produce reliable frequency counts and to subdivide these areas across various dimensions according to the varieties of language in which a word is used. By doing so the lexicographer is able to decide if an item has to be included or not in the standard dictionary.

6.4.2. Perspectives

Any modern lexicographer recognizes the need for compiling a corpus. In fact, most contemporary lexicographical works require the examination and analysis of collections of diverse kinds of spoken and/or written texts, commonly called corpora. In many publications on corpora, terms such as *representativeness* and *balance* are often associated with the compilation of "well-founded corpora". A *balanced corpus* contains texts from different genres and domains of use including spoken and written, private and public, etc. while a *representative corpus* covers the typical and central aspects of the language, and provides enough occurrences

of words and phrases. The real challenge for Fang lexicographers in the future will be the compilation of balanced and representative corpora in Fang.

The compilation of a corpus involves three main steps: 1) corpus design, 2) text collection, and 3) text encoding. The lexicographer has to go through all these steps (presented in chapter 3) if he/she wants to compile a dictionary that reflects the language of the target users. These steps can often be achieved with the assistance of computer software, but the use of this software has to be mastered by the lexicographer. To reach that objective, it is important that lexicographers should regularly have training in lexicomputing. In fact, despite the fact that a lexicographer can ask someone else to compile a corpus, it is better for him/her to know how to query a corpus in order to manipulate it, at any time, and then be able to retrieve what he/she wants from it.

In the recent past, corpora were mainly used for word frequency. Nowadays they are also used to assist the lexicographer in further, relevant aspects of language analysis such as collocations, concordances, etc. In the future, corpora could help lexicographers with grammatical analysis thanks to more sophisticated computer software.

As far as bilingual lexicography is concerned, the development of parallel corpora will help with the compilation of multilingual dictionaries and will play an important role in the improvement of automated translations. This means that huge corpora, e.g. that of COBUILD, should be translated in a maximum of languages in order to allow an easier compilation of translation dictionaries. In many

languages, particularly those with a strong oral tradition (such as the Gabonese languages), the compilation of huge corpora is not easy, since there are not too many written documents and the collection of important spoken material is a time consuming activity. The compilation of parallel corpora could help to make available important data in these languages. In Fang for instance, available corpora in English and French could be directly translated to Fang in order to ease the compilation of translation dictionaries between Fang and these languages. However, the use of parallel corpora does not mean that the compilation of corpora based on data retrieved directly from Fang should be abandoned.

6.5. Aspects of the dictionary structure

6.5.1. Review

Chapter 4, *The Aspect of the dictionary structure*, deals with different partial structures of the dictionary and the distribution of data in each dictionary structure. That distribution allows the lexicographer to determine the specific position of each data type in the dictionary. This chapter identifies the three main components of dictionaries: front matter, central list and back matter. Chapter 4 gives a detailed account of the data types that should be included in each of the three component parts of the proposed dictionary model. This account is preceded by a general overview of dictionary structures as proposed by Hausmann & Wiegand (1989) in order to better understand their approach and to apply it more easily to the proposed dictionary model. Hausmann & Wiegand's approach also constitutes the basis of the theoretical framework that is used to make provision for the different search zones within the dictionary articles. In this chapter, various specific points of the proposed

model are highlighted, and particular attention is paid to some parts of the central list, especially the macrostructure, the microstructure, the access structure, the addressing structure and the mediostructure.

As far as the macrostructure is concerned, chapter 4 describes that structure of the dictionary as the selection of lexical items to be included in the dictionary as lemma signs, the primary treatment units of the dictionary. It is the macrostructure that determines under which lemma a lexicographical item is to be found. In this chapter, the discussion about the macrostructure starts with synopsis of the macrostructure of three main dictionaries that exist in Fang. These dictionaries are the *Encyclopédie Pahouine* of Largeau (1901), the *Lexique fãn-français* of Martrou (1924) and the *Dictionnaire fang-français et français fang* of Galley (1964). Different types of macrostructures that can be found in dictionaries are also presented in that section. In fact, two types of macrostructures are identified; (1) a macrostructure with a straight alphabetical ordering, and (2) a macrostructure with a sinuous lemma file. The discussion of aspects related to the macrostructure is concluded with the suggestion of one specific type of macrostructure that should be suitable for the proposed dictionary model. The proposed macrostructure has to be in straight alphabetical ordering and should not include sublemmata. That position is motivated by the fact that the target users do not have a well-established dictionary culture and the inclusion of sublemmata could therefore confuse them. It is also important to note that in the Fang context, four main criteria should guide the selection of items to be included in the dictionary: (1) the frequency of the

items amongst the target group. (2) The function of the items, i.e. the role played by the items in the acquisition of knowledge in Fang or their importance in academic subjects or in the current environment. (3) The semantic value of the items, e.g. whether they are polysemous. (4) The status of the items as reflecting the culture of the target users.

The itinerary followed to dissert about the macrostructure is almost identical to the one used for the discussion about the microstructure in chapter 4. In fact, a more detailed account of the microstructure is also given in that section. That discussion starts with an overview of the microstructure in the three existing Fang dictionaries already used for the presentation of the macrostructure. It is followed by an inventory of the main types of microstructures that can be found in different types of dictionaries. From that inventory, three categories of microstructures are identified, i.e. unintegrated, integrated and semi-integrated microstructures. The unintegrated microstructure means that all semantic items, which are addressed at the lemma, appear in the first subcomment on semantics of the article, so that this is a comment on the lexical meaning. And all example items and items are distributed in the subcomment on semantics coming after the first one (cf. Gouws, 2001b). An integrated microstructure displays a system of direct addressing between a paraphrase of meaning/translation equivalent and its co-text entry/entries. Each paraphrase of meaning/translation equivalent is immediately followed by the co-text entry illustrating the typical usage of the lexical item in question (cf. Gouws, 1999b). A semi-integrated microstructure is a hybrid form displaying features of both an integrated and an

unintegrated microstructure and is typically used in more comprehensive dictionaries where lengthier articles with a variety of data types and search zones occur (cf. Gouws, 2001b).

The discussion of the microstructure continues with the presentation of the suggested microstructure for a French/Fang bilingual dictionary. Three main components are suggested for the microstructure of the proposed model: (1) a first obligatory component displaying the comment on form or what Hausmann & Wiegand call synchronic identification of the lemma (i.e. the phonetic pronunciation & the part of speech), and the comment on semantics of the lemma or the explanatory data of the lemma (i.e. paraphrase of meaning of the lemma in French). (2) A second obligatory component, considered as the most important, and entirely presented in Fang, displaying the equivalent item of the lemma, followed by the comment on form and the comment on semantics of that equivalent item. That component also displays illustrations and examples related to the equivalent item, and are complemented by some optional entries representing antonyms, synonyms, etc. related to the Fang equivalent item. (3) Lastly a third optional component that will depend on the treated Fang equivalent item. It will display additional explanatory features or cultural data related to the Fang equivalent item such as encyclopaedic or ethnological data. This component will be written in French because the intended target users are not yet able to read fluently in Fang.

Finally, the discussion of the microstructure gives a more detailed account of the types of equivalence that could be found in bilingual dictionaries and how some difficulties related to the

translation from the source to the target language could be solved. Three main types of equivalence are identified: absolute, partial and surrogate equivalence. (1) The absolute or congruent or full equivalence is when the meaning of a lexical unit in the source language is exactly the same as the meaning of another lexical unit in the target language. (2) The partial equivalence is when the meaning of source language items refers to one part or aspect of the meaning of the target language one (or vice versa). (3) The surrogate equivalence is when the target language has no lexical item to be used as translation equivalent of the lemma from the source language. In the case of surrogate equivalence in the compilation of a French-Fang bilingual dictionary, the use of three strategies is proposed: (1) the lexicographer can borrow words from other languages with the same semantic value and the same register of communication when they are already known and used by Fang people. (2) Substitutive terms can be created either according to the physical description of the object to be named, by comparing the object to be named with other objects, or according to the function of the object to be named. And (3) the most strongly recommended strategy will be for the lexicographer to conduct an investigation among speakers of the speech community and ask them how they can intuitively name the new object.

Chapter 4 also discusses other types of structures in a dictionary, i.e. mainly the access structure, the addressing structure, and the mediostructure. Regarding the access structure, chapter 4 highlights the fact that the success of a bilingual dictionary depends largely on the success that the typical target user achieves during a normal dictionary consultation procedure. In

this regard, the rapid and unimpeded access of the user to the relevant data presented in the dictionary has to be regarded as a prerequisite for successful lexicographic production in a user-driven approach (cf. Gouws, 2000a). The procedure of accessing a dictionary and following a search route is, in chapter 4, devised as part of the dictionary plan in the lexicographic process of the proposed dictionary. Two major types of access structure are discussed; (1) the outer access structure, which is the part of the search route, which leads the user from the entries on the cover of the dictionary to the lemma sign given as guiding element of the article, and (2) the inner access structure that starts at the lemma and proceeds through the dictionary article, to lead a user within an article to a required entry. The lemma sign is the final destination of the outer access structure and the starting point of the inner access structure.

For the proposed model, the access structure should be as precise as possible. That is why it is proposed that apart from the dictionary title, the cover should also specify the number of treated lemmata, the data categories included in the dictionary, the types of illustrations, etc. that should help the user to know what kind of data is available in the dictionary. A rapid access structure should also be available on each dictionary page, i.e. two running heads should be presented to indicate the first lemma at the left hand top and the last lemma at the right hand top of the page. In addition, the sections where outer texts will be included should be identified on the cover by different color stretches. Chapter 4 also advises the use of both typographical and non-typographical markers in the inner access structure for the proposed model. Concerning the

typographical structural markers, lemma signs and their translation equivalents should be presented in bold, paraphrases of meaning should be in normal roman characters and illustrative examples in italics. This presentation should allow the user who is only interested in finding a specific category of information to have a rapid access to the article slot where the data category he/she is looking for is accommodated. As far as the non-typographical structural markers are concerned, it is proposed not to utilize too many specific markers, because they can confuse users. For the proposed model only 4 symbols (cf. 4.2.5.1) should be given to introduce different data categories.

The new concept of the dictionary structure as proposed by Hausmann and Wiegand (1989), and identified in chapter 4, stipulates that the microstructure of the dictionary articles is not the only (partial) structure within the complete article structure. The other important structure is the addressing structure, which is the system according to which procedures of one entry directed at another is employed, i.e. each item refers to an addressee through an address given in the article. Normally the central address of a dictionary article is the item giving the form of the lemma sign and consequently the lemma. When the address is the lemma, one has a *lemmatic addressing*; when it is the sublemma, one has a *sublemmatic addressing*; in all other cases the items are addressed by *non-lemmatic addressing*. If all items inside the article are addressed at the lemma, we have *full lemmatic addressing*. Chapter 4 specifies that the microstructure of the proposed model should display a lemmatic addressing because some data such as pronunciation, the part of speech and the paraphrase of meaning

will be addressed at the lemma. However, that lemmatic procedure will only serve as guiding element that will lead the user to the main address, i.e. the equivalent item. That model should not display any sublemmatic addressing, since it will not include sublemmata. The main lexicographical treatment will be addressed at the equivalent item; this means that all the items inside the article are not exclusively addressed at the lemma. Consequently, the microstructure will not display a full lemmatic addressing. The proposed model should essentially utilize a non-lemmatic procedure.

The main lexicographic treatment will be addressed at the equivalent item and not at the lemma. Indeed, for a successful retrieval of the information needed by the user, a system of topic switching is necessary. Addressing procedures will not be directed at the macrostructural domain, but they will mainly involve the microstructure of the dictionary. A system of topic switching in the proposed model will essentially concern translation equivalents and all data categories addressed at these translation equivalents. Three component parts are proposed for the microstructure of the dictionary model: (1) an obligatory component that essentially displays the comment on form and some aspects of the comment on semantics of the lemma. (2) A second obligatory component (the most important), entirely in Fang, which displays in bold the equivalent item of the lemma, followed by its comment on form and comment on semantics. (3) A third component, which displays additional explanatory features or cultural data related to the Fang equivalent item such as encyclopaedic or ethnological data.

Chapter 4 also pays attention to the mediostructure, which is the system of reference employed in a dictionary to lead the user from one entry to another. It is specified that this reference procedure may be restricted to the specific article or it may exceed the boundaries of the article. Three types of reference markers are envisaged for the proposed model: (1) a double face arrow (\Leftrightarrow) referring the user toward the treatment of other lexical items where he/she can find additional data regarding the lemma or the equivalent item or any other entry in the article. (2) Abbreviations as *Syn.* (i.e. synonym), *Hom.* (i.e. homonym), *Ant.* (i.e. antonym), etc. referring the user to other lemma(s) or equivalent item (s) that have a semantic and/or morphological relation with the treated lemma or equivalent item. (3) Page numbers, also called double address entries, which involve the first two procedures, because the referred items should not be given only with the arrow or the mentioned abbreviations, but with the page number (in superscript) indicating where they can be found in the dictionary. In the mediostructure, two kinds of cross-references should be presented: the *article-internal cross-referencing*, which works within the boundaries of article and the *article-external cross-referencing*, which refers a user to an entry in another article or other text in the dictionary.

The last aspect discussed in chapter 4 concerns texts that contain an explanation of the editorial system as well as data added to complement the central list. This additional data is presented in the texts outside the central list either before the central list (*front matter texts*) or after the central list (*back matter texts*). It is specified that the use of outer texts should allow the lexicographer

to include entries that would not typically appear in a linguistic dictionary. A distinction is made between *integrated* and *unintegrated outer texts*. Unintegrated outer texts complement the central list and are not needed to retrieve the information presented in the articles of the central list. Integrated outer texts function in co-ordination with the central list and participate in achieving the genuine purpose of the dictionary. For the proposed model, the front matter should give an explanation of different structures of the dictionary particularly the article components as well as all symbols, icons, characters or signs used. That section should also give guiding notes on pronunciation, as well as abbreviations. The most frequently used verbs in Fang with their conjugation should be presented in the front matter texts. This presentation would represent a communication-directed function because the aim is to help the user with the production of texts in Fang. That section of the dictionary is also identified as the place where data should be provided about the genuine purpose of the dictionary, the intended user group, the general organization of the dictionary, the scope of the dictionary, etc. All these aspects should be explained in a way understandable to laypersons.

As far as the back matter is concerned, three main data categories are identified to be included in that section: (1) the first one should present Fang proper and common names with some encyclopedic data related to these names. (2) The second section should present pictures of certain domains or aspects that could not be given, in detail, in the central list. (3) The last section of the back matter should present some additional data, such as Fang dialects and their geographic location in Gabon and in Africa, etc.

Finally, chapter 4 is concluded with a discussion concerning the use of back matter texts for the treatment of synopsis articles. In fact, in a synopsis article the lexicographer treats the relevant lemma, but also gives data directed at the other lemma signs included elsewhere in the dictionary.

6.5.2. Perspectives

The structure of a dictionary is strongly related to its typology and purpose. The proposed dictionary model tries to adopt a structural organization, which is as simple as possible in order to allow easy consultation for the users. However, the proposed dictionary structure is also influenced by the data categories to be included in the microstructural treatment. As the proposed model is a monoscopal dictionary (from French to Fang), it should also be the basis that will support the compilation of the second section of the dictionary, as well as the compilation of future monolingual dictionaries in Fang. That is why it is suggested that the primary lexicographical treatment should be directed at the Fang equivalent items in the proposed model. When the reading and writing of Fang will become a common practice, as is the case with French, in Gabon, the second obligatory component of each article in the proposed microstructure could be retrieved to form the basis of a new monolingual dictionary.

As far as the second aspect of the scope of the dictionary is concerned, i.e. the Fang-French volume, a full lexicographical treatment of the Fang items should be maintained. Nevertheless, this treatment will not be addressed at the equivalent item, as in the first section, but at the lemma since Fang will be the source language in that section. Furthermore, the presentation of French

equivalent items in that dictionary should not display any comment on form or comment on semantics. In other words, after the treatment of lemmata, equivalent items should be given without any treatment. The motivation of that approach is that the proposed dictionary is intended to help with the standardization of Fang and not French, and there are already many French dictionaries where the user can find information about French items.

Nowadays, most metalexicographical publications are advocating a clear explanation of the dictionary structure in the user's guidelines in the front matter texts in order to help the user with an optimal retrieval of the needed information from the dictionary. However, many dictionary users are complaining about the fact that these guidelines do not always help them; either because the dictionary structure is too complicated and the explanation cannot be of any help, or the explanation is not explicit enough. In my view, despite the fact that the user's guidelines are important and must still be presented in modern dictionaries, the lexicographer should try to compile dictionaries with a very simple structure or organization so that the user can use the dictionary efficiently without going to the user's guidelines every time. In the Fang situation, a dictionary culture is not yet well established. Many users do not consult the user's guidelines. Therefore, the lexicographer must do things in such a way that users can easily find what they are looking for in the dictionary without necessarily consulting the user's guidelines. To reach this goal, Fang lexicographers should pay more attention to the dictionary micro-architecture that allows the user to easily identify the article slot or search zone where the needed information is located.

A dictionary structure includes many partial structures and each of them plays an important role in the accessibility of the dictionary. If the macrostructure, the microstructure, the addressing structure and the mediostructure cannot be the same in all dictionaries because they depend of the dictionary type, I think that some aspects on the *outer access structure* can be standardized. In fact, the outer access structure plays an important role when the user is looking for a specific dictionary or a specific lemma. That search starts from the dictionary cover and ends at the specific lemma sign inside the dictionary. Lexicographers can help the user to save time by systematically presenting a data exposure procedure on the dictionary cover, as well as the presentation of headers on dictionary pages in the central list. Thanks to the data exposure structure the lexicographer can inform the user about data categories that are included in the dictionary by naming them on the dictionary cover. The use of the outer access procedure by every lexicographer will also allow paper dictionaries to remain in competition with electronic dictionaries. Indeed, one of the biggest advantages of electronic dictionaries over paper dictionaries is their ability to allow users to quickly access the needed data thanks to their enhanced access structure. Paper dictionaries will be more attractive if their compilers improve and standardize the way they are presenting the outer access structure.

6.6. Dictionaries and standardization

6.6.1. Review

Dictionaries and standardization constitutes the fifth and last chapter of this doctoral dissertation. It is also the major discussion of the project because it demonstrates the relation between

dictionaries and the language standardization process. This chapter starts with the definition of the central concept, i.e. standardization, which is defined as *a collective term for those processes, which bring about uniformity in language by reducing diversity of usage or as the process, often in part at least deliberate, by which standard forms of a language are established.* In the specific case of a language, the standardization will be the process by which the choice of the standard variety will be made, as well as the promotion of that variety in the speech community. The definition of the concept highlights the fact that the standardization process is not applicable only on languages, but in any system that uses uniformity. Chapter 5 presents the general idea that standardization is *a process in which, among many ways of using or doing something (e.g. speaking a language), people will decide to recognize only one of these ways for general use in order to make the use easier and accessible for everyone, and consequently to improve the system involved.* The reasons that motivate the standardization of a language are also underlined in that section. It demonstrates that the development of a language involves its harmonization, i.e. the reduction of varieties in order to allow a better mutual understanding between speakers. Therefore, the standardization of a language is necessary to facilitate communication in a multilingual or multidialectal community. It also serves to establish an agreed orthography and provide a uniform form for school books. This will not only determine the form to be taught in language classrooms to native as well as non-native speakers, it will also decide the language of learning in general. The suppression of variation in language will lead to communication over

longer distances of space and time with a minimum of misunderstanding.

One of the important matters in the development of Gabonese languages is the lack of a specific alphabetical system that can allow the reading and the writing of these languages. The necessity to choose and use a unique writing system from those that exist is part of the discussion that is taking place in chapter 5. In this regard, apart from the historical overview of the concept of alphabet and an outline of the development of writing systems in Africa, chapter 5 also presents three alphabetical systems that can be used in writing of Gabonese languages: the Alphabet of Gabonese Idioms from Raponda Walker, the April 1999 alphabet from a workshop of experts, and the Rapidolangue Alphabet from the Raponda-Walker foundation. Particular attention is paid to the weak and strong points of each alphabet, and the quality that the standard alphabet should have is pointed out. Concerning that particular aspect, it is highlighted that there are no particular criteria for the choice of alphabetical symbols. It is a convention, and one can decide to use any kind of representation as alphabetical symbols. The only condition is that each sound of the language must correspond with a particular symbol in the alphabet, and that all relevant or pertinent sounds of the language are represented. In the same vein, it is underlined the acceptance, the teaching, the use and especially the adaptation are what make an alphabetical system strong, and all past, current or future proposals of the alphabets should be open for changes in order to improve them. In addition, it is specified that the choice of an orthographical alphabet based on the French one is not an "offence" against Gabonese languages. It is important

to readapt that alphabet according to the particularities of these languages. On the contrary, the familiarization of Gabonese people with the French orthographical system can be a good thing if the selected alphabet is close to French. The population will learn and understand it more easily. The matter concerning the alphabet is enlarged to a discussion about the use of tonal markers in the writing of Gabonese languages. Indeed, a disagreement exists between researchers concerning that matter. Some people estimate that the transcription of tones in writing of Gabonese languages is essential, even obligatory, since the latter play a relevant role in the functioning of these languages, others think that the transcription of tones is not useful for the users, particularly laypersons, because the good reading of tones is only accessible for people familiar with linguistic matters.

My own opinion is that the transcription of tonal markers in writing has failed in the majority of tone languages, however these languages are more developed than the Gabonese ones. Therefore, languages can be learnt without the transcription of supra-segmental markers. In addition, if Gabonese languages have to be included in the educational system, that means that they will be learnt at all levels of the education. In this case, it will be difficult to ask teachers to instruct students about tones when these ones (i.e. the teachers) do not know with certainty what it is all about (the teachers are not linguists). Finally, it has been proven that the more a language develops, the more its tonal system changes and it becomes simplified. In other words, languages are losing their tones as they evolve or are in contact with other linguistic systems. If tonal markers are included in the orthography of Gabonese

languages, they have to be regularly readapted every time there is any change in the language. I am against the use of tonal markers in the writing of Gabonese languages.

The matter regarding the writing of Gabonese languages is complemented by a discussion about the writing tradition. In fact, up to the present, two types of writing traditions have been used, i.e. the conjunctive and the disjunctive systems. Conjunctivism and disjunctivism are two different traditions of word division. The difference between these two systems concerns the status of certain linguistic elements. In conjunctivism, linguistic elements are joined to the following elements, whereas in the disjunctivism they are written separately.

As far as Fang is concerned, at the actual stage of the development of the language, it is difficult to take a definitive stand regarding the orthographical method of writing. In my view, the most important point is that the chosen written method must be the only one used by everyone for the orthographical writing of the language. The tradition of the users should also be taken into account. In fact, some people are already accustomed to writing some Fang words in a certain way. They do not care about the method these words are written, they just use them as they learnt them. Therefore, whatever the method retained, these words should be left as they are already used. However, I adhere more strongly to the disjunctive method and I recommend it for the proposed dictionary model. Writing words separately will allow the users to have a clear idea of the part of speech categories (e.g. noun, verb, adjective, subject, article, etc.) composing a sentence,

and consequently they can be able to use these items distinctly or in other contexts.

The most important discussion in chapter 5 is the one related to the relation between dictionaries and the language standardization process. That discussion starts with a preliminary analysis of what gives dictionaries authority to influence the standardization of languages. The term "authority" here refers to the power the dictionary can have amongst users by influencing their actions, opinions, or beliefs. In the Gabonese context, chapter 5 highlights the influence of the dictionary among four categories of users: (1) the speech communities, (2) the learners, (3) the educational system, and (4) the socio-professionals. In fact, the dictionary is considered by the speech community as the ideal source to improve its competence in language, when it is the main source of the linguistic competence of learners. The compilation of dictionaries in various domains of education, explains their increasing influence in educational fields. In the socio-professional field, many are using dictionaries to improve their knowledge not only in their own field, but also in various domains of life under the pressure of globalization. In that regard, dictionaries are very useful to them.

Chapter 5 identifies different practical ways in which dictionaries can really influence the development of Fang. Indeed, the role dictionaries can play in the standardization of Fang is not only achievable because of their authority among users, but also because of the role they can play in the preservation, and the development of the language. That is why that section emphasizes some of the important relations that can exist between a dictionary

and a language, in particular regarding aspects related to the conservation, the promotion and the development of Fang. Particular attention is paid to the relations between the dictionary and contemporary Fang and the reading and writing of Fang. The role the dictionary can play with regard to the use of Fang as academic subject and language of teaching, the acceptance of the standard dialect of Fang, the pronunciation of Fang, and the history of Fang are also emphasized.

Concerning dictionaries and contemporary Fang, chapter 5 highlights the fact that any dictionary compiled in any language at a specific period represents the language of that time. Even when the language is not spoken anymore it remains alive in the dictionary. Therefore, dictionaries can reflect the diachronic progression of Fang, and they can also show how Fang has been upgraded as a result of new epochs and events. One of the most important steps Fang needs to reach is to move from an oral language to a language of reading and writing. The dictionary can help it pass from one stage to another by promoting the standard writing, and giving users the habit to read and use the language through the dictionary. The need of young Fang people to learn their language is evident. Through their language, they want to save and develop their identity and their culture. Therefore, the dictionary can be an instrument in the teaching of Fang. Any modern dictionary includes and presents all aspects necessary for the learning and improvement of a language.

As far as standardization is concerned, chapter 5 identifies four types of dictionaries that can influence the standardization of Fang: (1) Dictionaries that aims at creating a written standard.

These so-called standard-creating dictionaries are usually the outcome of cultural activities in the speech community. For instance, the need to teach and introduce new concepts in Fang will show the necessity to write them in the language, i.e. bilingual dictionaries can be compiled in order to introduce into the standard form of Fang terms lacking in it but present in the source language. Standard-creating dictionaries can also be instruments in the survival of Fang by allowing it to introduce itself as a written language. Dictionaries of regional varieties of the standard Fang can help those varieties to develop a standard of their own through the compilation of small monolingual dictionaries. By compiling dictionaries with the standard Fang and its regional varieties, the dictionary will allow users to see the difference between standard Fang that has to be used in official and academic activities and the other varieties that can be used in any other contexts.

(2) Another type of dictionary that can develop the standard form of a language is the so-called modernizing dictionaries. As in any language, constant change of the lexicon is observable in Fang, because of the incessant influx of new words and expressions. Dictionaries of scientific and other terminology can record such new coinages. Their purpose is to define terms, which are consistently used in order to communicate effectively.

(3) Antiquating or archaising dictionaries can also influence the development of the standard Fang. This type of dictionary helps to use classical expressions instead of common language. They try to reverse changes in the language by re-introducing ancient and correct words. In the Fang situation, most young people are using borrowed words when they are speaking, because many Fang terms

and expressions are not used anymore, even when they exist. That situation can give the impression that the language does not have the appropriate words to name things, and has to borrow them from other languages. Antiquating dictionaries will help to bring back these words and re-introduce them in the language.

(4) Finally, standard-descriptive dictionaries aim to help the standard dialect to function as a generally and identically comprehensible means of communication beyond the boundaries and limits of the other language varieties. The standard dialect of Fang will need stability to remain generally and identically comprehensive for some period. Many users of a Fang standard dictionary will be interested in getting advice on how to correctly use the language particularly in educational fields. Such advice will be either implicit by excluding non-recommended items and usages from the dictionary or explicit by employing labels or usage notes. Chapter 5 also recommends that Fang standard-descriptive dictionaries should also give an important place to the contemporary language. This means that besides the standard-descriptive layer, these dictionaries should contain data about obsolete words not used anymore but found in literary works that are still read (e.g. dialectal expressions, colloquialisms, demotic speech, taboo words, etc.).

Chapter 5 also analyses means and methods by which the described dictionaries can influence the user, with a particular application to Fang. Indeed, among several synonyms or other variants, only one or two of them are listed in the dictionary. That means that the user who will try to communicate in the new standard Fang would use the listed rather than the non-listed

varieties. The selection of information (lexical units, variants of usage, etc.) is an influential factor because users do check the dictionary particularly when drafting written texts, and frequently accept what the dictionary says. Labels and symbols that give information about the status of expressions (e.g. colloquial, informal, dialect, obsolete, etc.) or that give the indications of syntactic patterns, etc. will also influence the way the user will utilize Fang. For instance if an expression is labeled as informal in the standard Fang dictionary, the user will probably try not to use it in formal situations (e.g. official speeches, academic writing, formal letters, etc.). Larger dictionaries contain short commentaries of various types within the article. These commentaries will influence the value of some words to the user. Dictionaries with historical slant for instance can and do give a historically founded argument for or against a usage. Dictionaries of technical terminology sometimes give encyclopaedic arguments to motivate the selection of a term instead of a competing synonym.

The discussion concerning the relationship between dictionaries and the language standardization process is complemented by an analysis of the place of electronic dictionaries in that development. In fact, the evolution of dictionaries in the future will probably be connected to that of computers. Computer systems offer natural language interfaces used in different environments and formats, and generally called electronic dictionaries. With globalization, multilateral exchanges, and perpetual progress of the technology, etc. electronic dictionaries will become accessible and more frequently used by the majority of dictionary users. That is why they can play an important role in the

standardization and growth of developing languages such as Fang. In that section, apart from the presentation of electronic dictionary typology, chapter 5 identifies advantages of electronic dictionaries, e.g. where the knowledge of the alphabet is needed for the efficient use of a paper dictionary, the user can access the wanted information in a CDROM dictionary by typing the first letters or words related to that information. Apart from the fact that the capacity of an electronic dictionary can largely go over 600 Mega octets (Mo), electronic dictionaries are easy to carry because of their weight and their format. In this regard, chapter 5 recommends the compilation of electronic dictionaries in Fang that could integrate games, and give users the possibility to combine business with pleasure, by learning and playing at the same time. Chapter 5 also pays particular attention to the role the Internet can play in the standardization process of Fang. Indeed, many on-line dictionaries and lexicographical reference works (e.g. automatic translations) are nowadays available on the Net. Any person having a computer connected to the Internet can access these on-line lexicographical works. The publication of Fang dictionaries on the Internet could be an excellent means to standardize and develop that language. In the future, the use of Internet for personal and academic research will probably increase even in countries where the Internet access is still difficult today. The availability of on-line Fang dictionaries will probably influence the use of that language and enlarge its accessibility to more users. The construction of websites entirely in Fang, with interactive features or pages, could also play a significant role in the standardization process of Fang. That will

force the people to read in Fang and accustom themselves with the language.

6.6.2. Perspectives

A dictionary whose objective it is to describe the standard dialect has to record the standard orthography of lexical items where there are competing forms. Lexicographers should therefore decide which forms are considered standard and which are not. In the Fang situation where there is not long written tradition, a standard dictionary has only one mission: to propagate the orthography that the lexicographer estimates to be the right to be used as the standard one. That objective can only be achieved if standard dictionaries are largely spread in order to make them reachable to the majority of people. Therefore, one of the biggest challenges of compilers of Fang dictionaries will be to compile as many dictionaries as possible and particularly to distribute them in areas where their influence will be effective (e.g. academic institutions).

At the current stage of the development of Fang, the majority of dictionaries that will be compiled in that language after the adoption of the standard dialect and writing system will all somehow have a role in the standardization process. Even if these dictionaries do not have the standardization as main objective, they should at least stick to standard dialect and writing. A dictionary, whatever its typology and function(s), is regarded as the authority on spelling, grammar, meaning and usage of a language (cf. Mdee, 1999), therefore all Fang dictionaries will be as important as the proposed model in the standardization process of Fang.

Standardization involves arbitrary decisions by language planners (e.g. choosing one dialect of the language as norm), and

two of these arbitrary but useful decisions are the simplification of the spelling and the adoption of only one form of words that have variants forms. The standard dictionary should be able to execute these responsibilities. However, if the dictionary includes words of other social and regional dialects, the lexicographer should make sure that these words are stated unambiguously and they should be marked in order to indicate their geographical or social status.

The contribution of dictionaries in the standardization process of languages all over the world cannot be denied. Nevertheless, that contribution has been efficient not only because of the authority of dictionaries among the speech community, but particularly because of the “easy access” to these dictionaries by people. By “easy access” I mean the way these dictionaries were compiled (e.g. orthography, presentation of data categories, micro-architecture, etc.), and their availability in areas (e.g. libraries, bookshops, schools, offices, etc.) that could allow the majority of people to consult them. In the Fang situation, two important issues should be solved first if one wants dictionaries to fully assist the standardization of that language. The first issue is the alphabetical system. Indeed, all Fang dictionaries have to be compiled in a unique writing system in order to accustom the user with that system. That will allow them to read the language fluently and identify Fang words wherever they are written. A dictionary cannot influence a user if he/she is unable to read it.

The second matter that has to be solved by Fang lexicographers is the access of existing or future Fang dictionaries to the majority of people. Until now it is almost impossible to find a Fang dictionary in a library or a bookshop in Gabon. Most existing

Fang dictionaries are conserved as museum pieces in Europe. Many people do not even know that they exist. In this regard, the compilers of future Fang dictionaries have to make sure that their products are really accessible to the majority of people, if they want these dictionaries to influence the use of Fang, and consequently its standardization. This means that Fang lexicographers should also make provision for a marketing policy in order to sale and/or distribute their products. For instance, an agreement could be reached with library managers and heads of academic institutions so that they can get Fang dictionaries for free and make them available to the public at large. Although, the lexicographer will receive some royalties according to the rate of consultation of these dictionaries by users.

6.7. General conclusion

Metalexicography or theory of lexicography is generally described as a whole range of activities concerned with the status of the field of lexicography (cf. Hartmann & James, 1998). Nowadays, it becomes a more practical activity primarily directed at the formulation of theoretical models to improve the quality of dictionaries (cf. Gouws, 2001a).

The present research project had two major objectives: (1) to formulate metalexicographical criteria for the compilation of a *French/Fang standard translation dictionary*, and (2) to highlight the important role that the proposed dictionary model could play in the standardization of Fang. The formulation of metalexicographical criteria for the compilation of the intended dictionary took place through chapters 3 & 4 where I presented: (a) the compilation process of a corpus that can serve for the compilation of a

French/Fang standard translation dictionary (cf. chapter 3), and (b) the organization or presentation of different partial structures of the intended dictionary (cf. chapter 4). The role that the proposed dictionary model can play in the standardization process of Fang is discussed in chapter 5. This discussion is a kind of demonstration that shows what types of features the proposed dictionary should have in order to influence the use of Fang in Gabon.

As any theory, the present metalexical project is not perfect. It can still be improved in many aspects. However, if the metalexical criteria formulated in this dissertation can assist any lexicographer during the compilation of any dictionary and that this dictionary is recognized as a good quality product, then I will be able to claim that I achieved my mission as metalexical.

BIBLIOGRAPHY

I. Dictionaries & Encyclopedias

- Galley, S.** 1964. *Dictionnaires Fang-Français et Français-Fang, suivi d'une grammaire Fang*. Neuchâtel: Henri Messeiller.
- Hartman R.R.K. & G. James.** 1998. *Dictionary of lexicography*. Routledge. New York.
- Largeau, V.** 1901. *Encyclopédie Pahouine. Elément de grammaire et Dictionnaire Français-Pahouin*. Paris: E. Leroux.
- Larouse.** 1971. *Nouveau Petit Larousse*. Paris: Librairie Larousse.
- Larouse.** 1998. *Petit Larousse Illustré*. Paris: Larousse.
- Lejeune, R.P.** 1872. *Dictionnaire Français-Fang ou Pahouin, précédé de quelques principes grammaticaux sur cette même langue*. Paris: A. Faivre & H. Teillard.
- Marling, R.P.** 1872. *Dictionnaire Fang-Français*. New York.
- Martrou, L.** 1924. *Lexique Fân-Français*. Paris Procure Générale (des Pères du St Esprit) / Abbeville: Imp. Paillard.
- Matthews, P.H.** 1997. *Concise Dictionary of Linguistics*. New York: Oxford University.
- Microsoft.** 2001. *Encarta Encyclopedia Standard*. Ireland. Microsoft.
- Sinclair, J.** (Ed.). 1995. *Collins COBUILD English Language Dictionary*. COBUILD: London: HarperCollins.
- Sinclair, J.** (Ed.). 2001. *CollinsCOBUILD English Dictionary for Advanced Learners*. London: HarperCollins.
- Webster.** 1995. *Merriam Webster Dictionary*. Zane Publishing. Dallas.

II. Primary sources

- Afane Otsaga, T.** 2000. *Esquisse phonologique et morphologique du Meka (langue bantoue de sigle A83)*. Unpublished, Mémoire de maîtrise. Libreville: Université Omar Bongo.
- Afane Otsaga, T.** 2001. Les dictionnaires bilingues dans les langues gabonaises: Approche méthodologique. *Eléments de Lexicographie Gabonaise, Tome 1*: 137-159. J. D. Emejulu (Ed.). New York: Jimacs-Hillman Publishers.

- Afane Otsaga, T.** 2002. Le rôle du dictionnaire dans le développement des langues gabonaises. *Éléments de Lexicographie Gabonaise, Tome II*: 206-229. J. D. Emejulu (Ed). New York: Jimacs-Hillman Publishers.
- Afane Otsaga, T.** 2002a. Les tons dans les dictionnaires des langues gabonaises: situation et perspectives. *Lexikos* 12:75-89.
- Afane Otsaga, T.** 2003. Ecriture et orthographe: critères, qualités, conditions & vulgarisation dans le cas du Gabon. *Ecriture et standardisation des langues gabonaises*. Unpublished. Libreville: Editions Raponda Walker.
- Al Kasimi, A.M.** 1977. *Linguistics and Bilingual Dictionaries*. Leiden: E.J. Brill.
- Alberts, M.** 2002. E-terminology. *Lexikos* 12: 90-104.
- Anthonissen, C. & D. Gough.** 1998. *A Pragmatic or Purist Approach to Language Standard: Implications for outcomes based educational*. Per Linguam. Vol 14. N1. Stellenbosch: University of Stellenbosch.
- Atkins, B.T.S. & A. Zampolli.** 1994. *Computational Approach to the Lexicon*: New York: Oxford University Press.
- Atkins, B.T.S, Beth L. & A. Zampolli.** 1994. Computational Approach to the Lexicon: An Overview. *Computational Approach to the Lexicon*: 17-37. Atkins, B.T.S. & A. Zampolli (Eds). New York: Oxford University Press.
- Atkins, S. Clear, J. & N. Ostler.** 1992. Corpus Design Criteria. *Literary and Linguistic Computing, Vol. 7, N 1*. Oxford:Oxford University Press.
- Barnhart, C.** 1980. What Makes a Dictionary Authoritative? : *Theory & Method in Lexicography: Western & non-Western Perspectives*:33-42. Columbia: Hornbeam Press.
- Béjoint, H.** 2000. *Modern lexicography. An introduction*. Oxford: Oxford University Press.
- Dubois, J. & C. Dubois.** 1971. *Introduction à la lexicographie: Le dictionnaire*. Paris: Librairie Larousse.
- Emejulu J. D. & Y. Nzang-Bie.** 1999. Linguistic Perspectives in Gabon. *Colloquium. Grand Forks*. University of North Dakota. SIL, Inc.
- Emejulu, J.D.** (Ed.). 2001. *Éléments de lexicographie gabonaise Tome 1*. New York: Jimacs-Hillman Publishers.
- Emejulu J. D.** 2001. Lexicographie Multilingue et Multisectorielle au Gabon: Planification, Stratégie et Enjeux. *Éléments de Lexicographie Gabonaise, Tome 1*: 38-57. J.D. Emejulu (Ed.). Jimacs-Hillman Publishers. New York.
- Emejulu J. D.** 2002. Défis et Promesses de la Lexicographie Intégrale dans les Pays en Développement. *Éléments de Lexicographie Gabonaise Tome 2*: 366-381. J.D. Emejulu (Ed). Jimacs-Hillman Publishers. New York.
- Emejulu, J.D. & Y. Nzang-Bié.** 1999. *Initial language development. Explaining Multilingual Education: Information on Some Tough Questions*: 39-65.

- Engwall, G.** 1994. Not Chance but Choice: Criteria in Corpus Creation. *Computational Approach to the Lexicon*: 49-82. Atkins, B.T.S. & A. Zampolli (Eds). New York: Oxford University Press.
- Gallardo, A.** 1980. Dictionaries and the Standardisation Process. *Theory and Method in lexicography*: 59-69. Zgusta, L. (Ed.). Columbia: Hornbeam Press
- Gouws, R. H.** 1989. *Leksikografie*. Pretoria/Kaapstad: Academica.
- Gouws, R. H.** 1996b. Bilingual Dictionaries and Communicative Equivalence for a Multilingual Society. *Lexikos* 6: 14-31
- Gouws, R. H.** 1997. Lexicographic Planning. *Lexicographic Consultative Meeting*: 16-22, PanSALB. Holiday Inn Johannesburg International Airport, 31 October 1997.
- Gouws, R.H.** 1999. Mediostructural Representation, Textual Condensation and User-Orientation in the WAT X. *Lexicographica* 15: 4-37. Tübingen: Max Niemeyer
- Gouws, R.H.** 2000. *Equivalent Relation in Translation Dictionaries*. M. Phil in Lexicography. Stellenbosch: University of Stellenbosch.
- Gouws, R.H.** 2000a. *General Theory of Lexicography: An introduction*. M. Phil in Lexicography. Stellenbosch: University of Stellenbosch.
- Gouws, R. H.** 2000b. Toward the Formulation of a Metalexigraphic Founded Model For the Lexicographic Units in South Africa. *Wörterbücher in Der Diskussion IV*: 109-133. Tübingen: Max Niemeyer Verlag.
- Gouws, R.H.** 2001. The Use of an Improved Access Structure in Dictionaries. *Lexikos* 11: 101-111.
- Gouws, R.H.** 2001a. Lexicographic Training: Approaches and Topics. *Éléments de lexicographie gabonaise Tome I*: 58-94. Emejulu J. D. (Ed.). New York: Jimacs-Hillman Publishers.
- Gouws, R. H.** 2002. Using a frame structure to accommodate cultural data. *Éléments de Lexicographie Gabonaise, Tome II*: 54-69. J.D. Emejulu (Ed.). New York: Jimacs-Hillman Publishers.
- Gouws, R.H.** 2002a. Niching as Macrostructural Procedure. *Lexikos* 12: 133-158
- Guthrie M.** 1953. *The Bantu Languages of Western Equatorial Africa*, Oxford: Oxford University Press.
- Hartmann, R.R.K.** (Ed.). 1979. *Dictionaries and their Users: Proceedings of the 1978 BAAL Seminar on Lexicography*, Exeter: University of Exeter.
- Hartmann, R.R.K.** (Ed.). 1984. *LEXeter '83 Proceedings. Papers from the International Conference on Lexicography at Exeter, 9-12 September 1983*, Lexicographica Series 1. Tübingen: Max Niemeyer.
- Hartmann, R.R.K.** (Ed.) 1990. *Lexicography in Africa. Progress Reports from the Dictionary Research Centre Workshop at Exeter, 24-26 March 1989*, Exeter Linguistic Studies Volume. 15, Exeter: University of Exeter Press.

- Hartmann, R.R.K.** 1992. Lexicography, with Particular Reference to English Learners' Dictionaries. *Language teaching*, 25:3, 151-159.
- Hausmann, F. J. & H. E. Wiegand.** 1989. Component Parts and Structures of General Monolingual Dictionaries: A Survey. Hausmann, F. J. et al. (Eds.). (1989-1991): 328-360, Berlin: Walter de Gruyter.
- Hausmann, F. J. & R. O. Werner.** 1991. Spezifische Bauteile und Strukturen zweisprachiger Wörterbücher: eine Übersicht. Hausmann, F. J. et al. (Eds.): (1989-1991): 2729-2769, Berlin: Walter de Gruyter.
- Hausmann, F. J.** 1981. Wörterbücher und Wortstchatzlernen. *Linguistik und Didaktik* 45/46: 71-78.
- Hausmann, F. J., Reichmann, O., Wiegand, H. E. & Zgusta, L.**(Eds.). 1989-1991. *Wörterbücher. Ein Internationales Handbüch zur Lexikographie/ Dictionaries. An International Encyclopedia of Lexicography/Dictionnaires. Encyclopédie Internationale de Lexicographie.* 3 volumes; 3355 pages, Berlin: Walter de Gruyter.
- Hausmann, F. J.** 1990. De quoi se compose l'article du dictionnaire de langue? L'importance du sous-adressage. T. Maguy and J. Zigány (Eds.): *BudaLex '88 Proceedings*: 59-66. Papers from the EURALEX Third International Congress, Budapest: Akadémiai Kiadó.
- Hombert, J. M. & A.-M. Mortier.** 1990. Bibliographie des langues du Gabon. *Revue gabonaise des Sciences de l'Homme* 2 : 335-355. Mise à jour de la bibliographie publiée dans Pholia. *Revue du Laboratoire de Phonétique et Linguistique Africaine* 1. 1984.
- Hombert, J. M.** 1990a. Les langues du Gabon: État des connaissances. *Revue gabonaise des Sciences de l'Homme* 2 : 29-36.
- Hombert, J. M.** 1990b. Atlas Linguistique du Gabon. *Revue gabonaise des Sciences de l'Homme* 2: 37-42.
- Householder, F.W. & S. Saporta.** 1962. *Problems in Lexicography*. Bloomington: Indiana University Press.
- Kachru, B.** 1980. The New Englishes and Old Dictionaries: Directions in Lexicographical Research on Non-native Varieties of English. *Theory & Method in Lexicography: Western & non-Western Perspectives*. Columbia: Hornbeam Press.
- Kromann, H.-P., T. Riiber, & P. Rosbach.** 1991. *Grammatical Constructions in Bilingual Dictionaries*. F. J. Hausmann et al. (Eds.), Volume III. 2770-5.
- Landau, S.I.** 1984. *Dictionaries. The art and craft of lexicography*. New York: C. Scribner's Sons/Cambridge: O.U.P.
- Landau, S.I.** 2001. *Dictionaries. The art and craft of lexicography*. Second Edition. Cambridge: Cambridge University Press.

-
- Lazore, D. K.** 1993. The Assembly of First Nations in its 1990 report: Towards Linguistic Justice for First Nations. *Conference Report, August 9-10*.
- Lodge, A.** 1993. *French: from dialect to standard*. London and New York: Routledge.
- Louw, P. A.** 1999. Access Structure in a Standard Translation Dictionary. *Lexikos 9*: 108-118.
- Mabika Mbokou, L.** 2001. Le dictionnaire et le système éducatif. *Éléments de Lexicographie Gabonaise, Tome I*:206-222 J. D. Emejulu (Ed). New York: Jimacs-Hillman Publishers.
- Malkiel, Y.** 1980. The Lexicographer as a Mediator Between Linguistics and Society. *Theory & Method in Lexicography: Western & non-Western Perspectives*: 43-58. Zgusta, L (Ed). Columbia: Hornbeam Press.
- Mavoungou, P.A.** 2002. Synopsis Articles in the Planning of a Trilingual Dictionary: Yilumbu-English-French. *Lexikos 12*: 181-200.
- Mavoungou, P.A.** 2001. La Mondialisation et la Lexicographie Trilingue ou Plurilingue au Gabon. Emejulu J. D. (Éd.). 2001: *Éléments de lexicographie gabonaise Tome I*: 160-183. New York: Jimacs-Hillman Publishers.
- Mavoungou, P.A.** 2001a. Macro- and Microstructural Issues in Mazuna Lexicography. *Lexikos 11*: 122-138.
- Mavoungou, P.A.** 2002. *Metalexical criteria for the compilation of a trilingual dictionary: Yilumbu-English-French*. Unpublished DLitt. dissertation. Stellenbosch: University of Stellenbosch.
- Mavoungou, P.A., T. Afane Otsaga & G.-R. Mihindou.** 2003. Dictionaries Compiled with French and the Reproduction of the Gabonese Cultures. *Lexikos 13*:133-153.
- Mdee, J.S.** 1999. Dictionaries & the standardization of spelling in Swahili. *Lexikos 9*: 119-134
- Nzang-Bie. Y.** 2002. Le corpus lexicographique dans les langues à tradition orale: le cas du dialect fang-mekè. *Lexikos 12*: 111-138
- Ondo-Mebiame, P.** 1992. *De la phonologie à la morphologie du Fang parlé à Aboumezok (Langue Bantu A.78)*. 2 volumes. Unpublished Thèse de doctorat. Tervuren: Université Libre de Bruxelles.
- Prinsloo, D.J.** 2000. *Electronic corpora as a basis for the compilation of African-language dictionaries. Part 1: The macrostructure*. Pretoria: University of Pretoria.
- Prinsloo, D.J.** 2000a. *The Compilation of Electronic Corpora, with Special Reference to African Languages*. Pretoria: University of Pretoria.
- Prinsloo, D.J.** 2001. The Compilation of Electronic Dictionaries for African Languages. *Lexikos 11*:139-159.
-

- Rey-Debove, J.** 1971. *Etude linguistique et sémiotique des Dictionnaires français contemporains*. Thomas A. Sebeok (Ed). The Hague: Mouton.
- Rey-Debove, J.** 1989. L'équivalence dans le dictionnaire bilingue. *Wörterbücher. Dictionnaires. Dictionnaires*: 2817-2823. F.J. Hausmann & al. (Ed). Berlin: De Gruyter.
- Sadembouo, E.** 1991. Préalable à la Standardisation des Langues Africaines: *Standardisation des Langues Africaines*. Cyffer, N. et al. (Eds.). Hamburg: Helmut Buske Verlag.
- Sadembouo, E., M. Tadadjeu & U. Wiesmann.** 1988. *Guide pour le développement des systèmes d'écriture des langues africaines*. Yaoundé: Université de Yaoundé & SIL-UdY.
- Svensén, B.** 1993. *Practical Lexicography. Principles and Methods of Dictionary-making*. Oxford/New York: Oxford University Press.
- Touré, A.** 1990. L'écriture des langues africaines: évolution et principes méthodologiques. *Revue Gabonaise des Sciences de l'Homme N 2*. Libreville: LUTO/Université Omar Bongo.
- Van Wyk, E.B.** 1995. Linguistic Assumptions and Lexicographic Traditions in the African Languages . *Lexikos 5*: 82-96.
- Wells, R.S.** 1994. *Dictionaries & The Authoritarian Tradition: A study in English Usage & Lexicography*. The Hague-Paris: Mouton
- Wiegand, H. E.** 1998. *Wörterbuchforschung*. Berlin: De Gruyter.
- Zgusta, L.** (Ed.). 1980. *Theory and Method in Lexicography*. Columbia: Hornbeam Press.
- Zgusta, L.** 1971. *Manuel of Lexicography*. The Hague: Mouton.
- Zgusta, L.** 1987. Translational Equivalence in a Bilingual Dictionary. *Journal of the Dictionary Society of North America 9*: 1-43, Indiana.
- Zgusta, L.** 1989. The Role of Dictionaries in the Genesis and Development of the Standard. Hausmann, F. J. et al. (Eds.). 1989-1991:
- Zgusta, L.** 1992. *History, Languages and Lexicographers*. Tübingen: Max Niemeyer.
- III. Secondary sources** (complementary literature not necessarily quoted in the dissertation)
- Andeme Allogho, M.F.** 1980. *Esquisse phonologique du Nzaman de Makokou*. Unpublished, Mémoire de Licence. Libreville: Université Omar Bongo.
- Bancel, N.** 1999. *Entre acculturation et révolution. Mouvement de jeunesse et sports dans l'évolution politique et institutionnelle de l'AOF 1945-1960*. Unpublished, Thèse de doctorat nouveau régime en Histoire de l'Afrique. Paris: Université Paris I-Sorbonne.

- Barreteau, D.** (Ed.). 1978. *Inventaire des études linguistiques sur les pays de l'Afrique Noire d'expression française et sur Madagascar*, Paris: CILF.
- Bergenholtz, H., & S. Tarp.** 1995. *Manual of Specialised Lexicography*, Amsterdam: John Benjamins.
- Bibang Meye.** 2001. *Etude des compétences narratives chez les enfants bilingues Fang-Ntumu-Français d'Oyem (Psycholinguistique): Etude transversale sur le développement de la temporalité et de la connectivité chez des sujets âgés de 7 à 12 ans*. Unpublished Mémoire de Maîtrise. Libreville: Université Omar Bongo.
- Carpentier de Changy, H. & Voltz, M.** 1990. Alphabet scientifique des langues du Gabon: liste alphabétique. *Revue Gabonaise des Sciences l'Homme* N 2: 113-115. Libreville: LUTO/Université Omar Bongo.
- CICIBA.** 1997. *Rapport du Séminaire sur la lexicographie Bantu*, Rapport non publié, Libreville: CICIBA.
- Cinnamon, J.** 1990. Fang (nzaman). *Revue Gabonaise des Sciences de l'Homme* 2 : 175-183. Libreville: LUTO/Publications de l'Université Omar Bongo.
- Doneux, J. L.** 1967. *Questionnaires d'enquête linguistique* (Greenberg-Tervuren-Welmers), Dakar: Université de Dakar.
- Ella Ella.** 2000. *Etude contrastive des pronoms relatifs en Français et en Ntumu*. Unpublished Mémoire de Maîtrise. Université Omar Bongo. Libreville.
- Ella, M.** 2000. *Etude linguistique d'une oeuvre littéraire: Un mvvet de Zué Nguema, chant épique Fang recueilli par herbert Pepper*. Unpublished Mémoire de Maîtrise. Libreville: Université Omar Bongo.
- Emejulu, J.D.** 2000. Lexicography, an Economic Asset in Multilingual Gabon. *Revue Gabonaise des Sciences du Langage/Gabonese Journal of the Language Sciences* 1: 51-69.
- Fondation Raponda-Walker.** 1995. *Rapidolangue: Méthode d'apprentissage des langues nationales*. Volume N°1 (fang, inzebi, lembaama, omyene, yipunu). Libreville: Editions Raponda-Walker.
- Fondation Raponda-Walker.** 1997. *Rapidolangue: Méthode d'apprentissage des langues nationales*. Volume N°2 (ikota, ghetsogo). Libreville: Editions Raponda-Walker.
- Galley, S.** 1889. *Nteni osi nteni Fañwe*, Talagouga: Imprimerie de la Mission.
- Gouws, R. H. & Prinsloo, D. J.** 1997. Lemmatisation of Adjectives in Sepedi. *Lexikos* 7 : 45-57.
- Gouws, R. H. & Prinsloo, D. J.** 1998. Crosss-Referencing as a Lexicographic Device. *Lexikos* 8 : 17-36.
- Gouws, R. H. & F. A. Povelis.** 1992. The Development of Afrikaans and the Lexicographical Tradition. Zgusta, L. (Eds.): *History, Languages and Lexicographers*: 77-164. Tübingen: Max Niemeyer.

- Guthrie, M.** 1953. *The Bantu Languages of Western Equatorial Africa*. Oxford: Oxford University Press.
- Hombert, J.M.** 1990. Problèmes phonétiques et phonologiques rencontrés dans les langues du Gabon. *Revue Gabonaise des Sciences de l'Homme N 2*. Libreville: LUTO. Université Omar Bongo.
- Hombert, J.M.** 1991. Quelques critères de classification des parlers fang. *Pholia 6*. Lyon: CNRS. Université Lumière-Lyon2.
- Idiata, D. F.** 2002. *Il était une fois les langues gabonaises*. Libreville: Editions Raponda Walker.
- Jacquot, A.** 1978. *Le Gabon*. Daniel Barreteau (dir), CILF: Inventaire des études linguistiques sur les d'Afrique noire d'expression française et sur Madagascar. Paris: SELAF.
- Kavanagh, K.** 2000. Words in a Cultural Context, *Lexikos 10*: 99-118.
- Kwenzi-Mikala, J.T.** 1988. L'identification des unité-langues bantu gabonaises et leur classification internes. *Muntu 8*: 54-64.
- Kwenzi-Mikala, J.T.** 1998. Parlers du Gabon. A. Raponda-Walker (Ed.): *Les langues du Gabon*: 217. Libreville: Édition Raponda-Walker.
- Lombard, F.J.** 1994. Lexicographer, Linguist and Dictionary User: An Uneasy Triangle? *Lexikos 4*: 204-214.
- Mayer, R.** 1990. Histoire de l'écriture des langues du Gabon. *Revue Gabonaise des Sciences de l'Homme N 2*. Libreville: LUTO/Université Omar Bongo.
- Mba-Nkoghe, J.** 1981. *Phonologie et classes nominales en fang* (Langue bantoue de la zone A). Unpublished Thèse de 3^{ième} cycle. Paris: Université Sorbonne Nouvelle.
- Mba-Nzué, N.** 1981. *Esquisse phonologique du Mvaï (Parler de Minvoul)*. Mémoire de Licence. Unpublished. Libreville: Université Omar Bongo.
- Medjo-Mvé, P.** 1997. *Essai sur la phonologie panchronique des parlers fang du Gabon et ses implications historiques*. Unpublished. Thèse de Doctorat (NR). Lyon: Université Lumière-Lyon2.
- Mekui Missang.** 1998. *La présence des faits suprasegmentaux du fang ntumu dans l'utilisation du français pour les enfants âgés de 10 à 12 ans*. Unpublished, Mémoire de Maîtrise. Libreville: Université Omar Bongo.
- Mouguiama-Daouda, P.** 1995. *Les dénominations ethnoichtyologiques chez les bantous du Gabon*. Thèse de doctorat, Lyon: Université Lyon 2.
- Mvé Ondo, B.** 1991. *Sagesse et Initiation à travers les contes, mythes et légendes fang*. Libreville-Paris. Centre Culturel Français Saint-Exupéry – Sépia.
- Ndinga-Koumba-Binza, H.S.** 2003. *Ecriture et standardisation des langues gabonaises: Approche historique et cas des langues gabonaises*. *Ecriture et standardisation des langues gabonaises*. Unpublished. Libreville: Editions Raponda Walker.

-
- Raponda-Walker, A.** 1932. Alphabet des idiomes gabonais. *Journal de la Société des Africanistes* 3(2): 305-314.
- Rey-Debove, J.** 1971. *Étude linguistique et sémiotique des dictionnaires français contemporains*. The Hague: Mouton.
- Smit, M.** 1996. *Wiegand's Metalexigraphy as a Framework for a Multilingual, multicultural, Explanatory Music Dictionary for South Africa*. Unpublished D.Litt. thesis. Stellenbosch: University of Stellenbosch.
- Soami, S.L.** 2003. La place des tons dans l'orthographe des langues gabonaises. *Écriture et standardisation des langues gabonaises*. Unpublished. Libreville: Editions Raponda Walker.
- Wiesemann, Sadembouo & Tadadjeu.** 1983-1988-2000. *Guide pour le développement des systèmes d'écriture des langues africaines*. Yaoundé: Coll. Propelca.

Index

| | Pages |
|---|--------------|
| Tables | |
| Table 1: Gabonese language forms classification | 10 |
| Table 2: Fang dialects locations | 12 |
| Table 3: Group Yaoundé-Fang | 51 |
| Table 4: Criteria, role and applicable fields | 76 |
| Table 5: Correspondence criteria/dialects | 76 |
| Table 6: Alphabet of Gabonese Idioms from Raponda Walker | 234 |
| Table 7: The April 1999 alphabet | 235 |
| Table 8: The Rapidolangue alphabet | 236 |
| Table 9: Examples of words with tones in some Gabonese languages | 245 |
| Schemas | |
| Scerba's classification | 24 |
| Malkiel's classification | 25 |
| Al-Kasimi's classification | 28 |
| Zgusta's classification | 31 |
| Hartmann and James' classification | 32 |
| Béjoint's classification | 35 |
| Afane Otsaga's classification | 38 |
| Schema of data collection | 109 |
| Schema of corpus updating | 121 |
| Global view of the dictionary structure | 127 |
| Macrostructure with a straight-alphabetical order | 137 |
| Macrostructure with a Niching ordering | 138 |
| Macrostructure with a Nesting ordering | 139 |
| Unintegrated microstructure in a monolingual dictionary | 165 |
| Unintegrated microstructure in a bilingual dictionary | 166 |
| Integrated microstructure in a monolingual dictionary | 168 |
| Integrated microstructure in a bilingual dictionary | 169 |
| Microstructure with a semi-integrated component in a monolingual dictionary | 171 |
| Microstructure with a semi-integrated component in a bilingual dictionary | 172 |
| Proposed microstructure | 177 |
| Figures | |
| Figure 1: Relationship between language and dialects | 56 |
| Figure 2: Global view of the standardization process | 221 |
| Figure 3: Development of Sumerian into cuneiform | 227 |

| | |
|--|-----|
| Index | 335 |
| <hr/> | |
| Figure 4: Development of Egyptian writing | 227 |
| Figure 5: Development of five letter from Egyptian hieroglyphs | 227 |
| Figure 6: Example of texts in Hausa | 250 |
| Figure 7: Example of texts in Kiswahili | 251 |
| | |
| Maps | |
| Map 1: Fang dialects distribution in Gabon | 13 |
| Map 2: Fang dialects distribution in Gabon | 68 |
| | |
| Pictures | |
| Examples of pictures in the back matter texts | 201 |
| Example of outer access structure on a dictionary front cover | 208 |
| Example of outer access structure on a dictionary back cover | 209 |
| Example of outer access structure on a dictionary spine cover | 209 |
| Example of outer access structure on a dictionary page | 210 |