Universität Bielefeld
Fakultät für Linguistik und Literaturwissenschaft

Learning Ega

A linguistic Frame of Reference for Primers for Endangered Languages

Magisterarbeit im Fach Anglistik
vorgelegt von Sophie Salffner

Bielefeld, März 2004

Gutachter
Prof. Dr. Dafydd Gibbon
Prof. Dr. Werner Kummer

CONTENTS

TABLES

GRAPHICS

INTRODUCTION

Universal Declaration of Human Rights

Article 2.

Everyone is entitled to all the rights and freedoms set forth in this Declaration, without distinction of any kind, such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status. Furthermore, no distinction shall be made on the basis of the political, jurisdictional or international status of the country or territory to which a person belongs, whether it be independent, trust, non-self-governing or under any other limitation of sovereignty.

Article 26.

(1) Everyone has the right to education. Education shall be free, at least in the elementary and fundamental stages. Elementary education shall be compulsory. […]

(United Nations 1948)

Article 2 of the Declaration of Human Rights guarantees the right of any person to one's own language (Ouane 2002, 72), Article 26 (1) guarantees the right to education. Taken together, these two articles may be interpreted the way that everybody is entitled to education in their own language. Whereas mother-tongue education has been made possible for official languages and languages with many speakers who have the financial means to pay for this education, things look nowhere near as good for languages with a smaller number of speakers and/or with less financial power. In many African countries for example, the language for education is the former colonial (and now national) language, no matter whether the children who start school speak the language or not.

This thesis will not discuss moral, ethical or political reasons for or against mother-tongue education. Instead, it works on basic assumption that mother-tongue education is both possible and necessary. The goal of the thesis is to set up a linguistic frame of reference for teaching material for mother-tongue education of spoken endangered minority languages and to apply aspects of this framework to a specific language so that teaching material in the form of primers can be designed which help the learners to learn or study their language.

In general terms, a primer gives an introduction to a given subject and is intended for beginners in the given field. Primers for spoken minority languages which are endangered have to meet different needs than "ordinary" primers because of the nature of the languages.

When a language is spoken (and not written), endangered (and not stable) and spoken by a minority (and not a majority), there are specific constraints on the primer which affect its content, structure and rendering and set it apart from primers designed for languages like English, which are both written and spoken, not endangered, used by a majority and are used as official languages. In addition, endangered languages are often spoken in developing countries, which puts additional, e.g. financial or infrastructural, constraints on the primer design.

Therefore, the general problem the thesis discusses is how an introductory textbook such as a primer for an endangered spoken minority language be designed with respect to content, structure and rendering so that it reflects and systematically takes into account the specific constraints.

I have not come across any research which looked at teaching endangered languages from a holistic point of view. No initiative I found covered all levels of linguistic analysis starting from pragmatics going down to phonetics and phonology. No paper or book I found gave linguistically based reasons for their choice of contents. All initiatives lack a well-founded linguistic model as a basis which puts content, structure, rendering and the constraints on these three aspects in relation.

Instead, the papers I read all treated specific aspects of teaching. Some researchers and organisations have a strong interest in developing literacy and practical orthography primers (e.g. SIL, Wycliff). Some put an emphasis on the ethnological and linguistic context and deduce possible settings for teaching and teaching methodologies from this (e.g. Flores Farfán 2001). Others discuss possible media for teaching material (e.g. Poser 2002) and again others describe organisational problems involved in teaching endangered languages in for example language workshops (e.g. Aikhenvald 2003).

Altogether, the literature leaves the impression that researchers work separately, on their own and on one language. I did not find any network or organisation which brings researchers together and allows them to exchange ideas. Also, I did not find any official platform such as a journal which specifically focuses on teaching and maintaining endangered languages. The only organisations I am aware of which work on a series of languages and do exchange ideas are the Christian organisations SIL and Wycliff. It is

symptomatic for the isolation in this field of research that I only found out that the second supervisor of this thesis has also already worked on primers for small languages when he told me himself – and that is within one university.
In a word, it is urgent and necessary to start working on a holistic systematic linguistic approach to primers for endangered languages which covers more than just one aspect and one language and to introduce the general model to other researchers for discussion.

This thesis starts out with a description of the language Ega and the Ega linguistic community, which will be focused on throughout the thesis. The first chapter sketches linguistic characteristics of the language, its sociolinguistic situation and its endangerment status. It outlines potential future prospects for the language relating to language learning, relearning, studying and language maintenance.
Chapter 2 explores the links between language documentation, teaching and primer design. It introduces the field of documentary linguistics and issues currently discussed there. Additionally, it shows how these issues apply to primer design.
In chapter 3, the concepts of document content, structure and layout and constraints on the document are introduced and combined in a linguistic model which is used as a basis for a primer. The intricacy with a model like this is that it has to be general enough to serve as a basis for a whole range of languages and, at the same time, it has to be specific enough to do justice to the idiosyncrasies of a particular language. The model presented here is intended to meet both demands. Also, the model can be understood as a meta paradigm which explains the rationale behind the structure of the following chapters 4 to 8.
Chapter 4 discusses the constraints a primer has to fulfil. Like with the model, both general constraints and the language-specific situation have to be taken into account. This is done by giving general constraints and applying them to Ega.
Chapter 5 discusses the question how the content for a primer can be set, what must be included and why. It covers the linguistic fields pragmatics, semantics, syntax, morphology, phonetics and phonology and orthography. Finally, a synthesised model of approaches is presented which is intended to guarantee the quality of the primer and its general linguistic and language-specific comprehensiveness.
Chapter 6 applies the general principles set up in chapter 5 to actual Ega language data and gives examples for Ega pragmatic issues, vocabulary, syntax etc.

Chapters 7 and 8 outline the proposals for structure and layout of a primer respectively. As this paper focuses on linguistic content more than on structure or layout of the primer, chapters 7 and 8 are considerably less detailed than the primer content chapters. Still, structure and layout are addressed and set in relation to the constraints on the primer defined in chapter 4.

To end with, chapter 9 takes up three issues discussed in the content of the primer and applies them to actual proposals for Ega teaching contents and material. The examples chosen for this application are an Ega orthography, Ega greetings and Ega story telling.

1. INTRODUCING EGA

1.1. *The Ega language*

1.1.1. Ega – homing in

Ega is a term for both an ethnic group living and a language spoken in the south-central region of the Ivory Coast in the Divo Department.

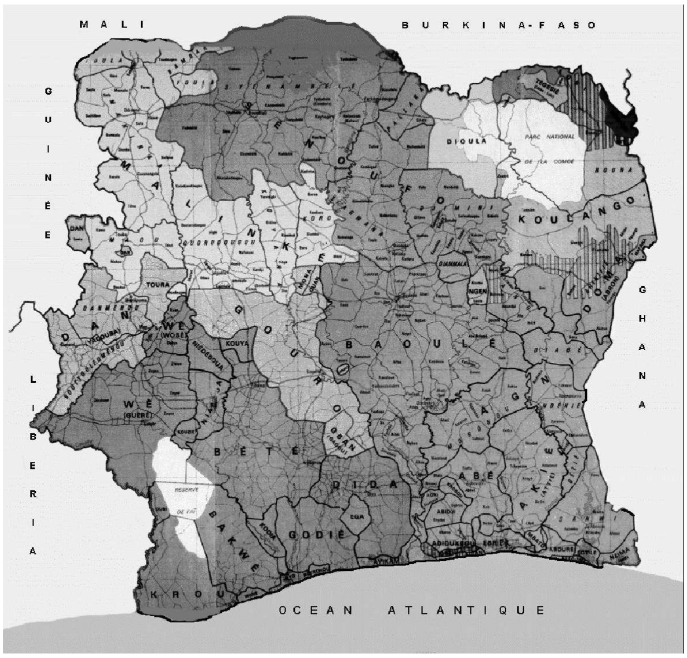


Figure 1: Map of the Ivory Coast

The language is called *Ega* /I-ga´/[1] by most of its speakers and most scientists and *Dies* by the Dida, another ethnic group living on that territory, and the Western Ega. The official administrative usage is also *Dies*; the Ethnologue lists *Egwa* as another possible name for the language. (SIL 2002)

The Ega live in twenty-one villages in three groups (North, Central, South). There are varying numbers concerning how many Ega there are. The 1963 census puts their number at 4300, the 1975 census at 67, Bole-Richard (1982, 359) estimates 5000 people, the Ethnologue (SIL 2002) gives a number of 291 to 3000. The current estimate is at 2000 people spread over 21/20 villages. (Ahoua/Connell/Gibbon 2002)

---

[1] All phonemic transcripts are given in Ega transcription conventions which use standard and simplified

Scientific interest and the fact that the language is considered to be endangered have led to its linguistic documentation. By 1999, the language had been fragmentally documented, since then more research has been carried out within the framework of the DOBES project[2]. The DOBES Ega project finished in 2002, efforts in documentation, however, remain. The current status is that information has been gathered relating to the documentation of Ega phonetics and phonology, morphology, lexicon, orature, sociolinguistic situation, cultural artefacts and Ega traditions. In addition, descriptive work has been carried out investigating Ega phonetics and phonology, morphology and Ega oral traditions. Also, there are various attempts at compiling Ega dictionaries.
(Connell/Ahoua/Gibbon 2002, 99; Ahoua/Connell/Gibbon 2002; Gibbon 1999, 6)

1.1.2. Genetic affiliation

The Ega language has been classified as belonging to the Niger-Congo languages and has been assigned to the Kwa family as the only member of the Nyo cluster. It is the most western of the Kwa languages and the only Kwa language spoken west of the Bandama River. There are no known closely related languages and no known related dialects. Ega is surrounded by Kru languages, namely Dida in the north and Godié in the south.

The classification of Ega as Kwa, however, is not a unanimous issue. According to Connell/Ahoua/Gibbon (2002, 99), this classification is based on only a small amount of comparative evidence and there are arguments against this classification.

First of all, Ega is structurally considerably different from any other Kwa language and has more complex phonetics, phonology and morphology than any other language of the Kwa group. The linguistic features of Ega will be illustrated in more detail in the next section. On the one hand, these differences may be taken as an indication that Ega is not Kwa. On the other hand, there is agreement among researchers that these linguistic differences may reflect archaic stages of Kwa language development. (Connell/Ahoua/Gibbon 2002, 99)

Secondly, there is strong influence of Kru languages (Bete and Godié), which may cast doubt on the classification of Ega as Kwa and may relate Ega more closely to Kru languages. Yet, Ahoua/Connell/Gibbon (2002) state that in terms of vocabulary only

---

X-SAMPA format, cf. http://coral.lili.uni-bielefeld.de/langdoc/EGA/Formats/egasampa.txt
2 URL: http://www.mpi.nl/DOBES/teams/Ega/Ega.html

10% of the Ega vocabulary is Kru, that no core vocabulary is included in this set and that the Kru vocabulary in the Ega language may be attributed to loans. Moreover, 30% of the vocabulary of Ega is Kwa, which – if the languages are related – hints more towards Ega being related to Kwa languages, even if it is only at a historical distance. In addition to the comparatively few lexical similarities, there are structural differences between the Kru languages and Ega. Ega uses noun prefixes whereas these do not occur in Kru languages. With respect to phonology, Kru languages do not have implosives or nasal vowels where Ega features both. (Ahoua/Connell/Gibbon 2002) Taken together, the evidence indicates that even if Ega is not Kwa, it is even more unlikely to be Kru.

Apart from the classification of Ega as a Nyo cluster isolate within the Kwa languages, there is a recent classification of Ega as an isolate within Kwa, co-ordinate with all other branches of Kwa. Finally, there is a minority opinion that Ega is not Kwa but a true Niger-Congo isolate. (Ahoua/Connell/Gibbon 2000, 1)

Description of Ega in this thesis is based on the working hypothesis that Ega is a Kwa language. This is because this classification seems to be the most widely accepted and because it is better backed up by evidence than the other classifications.

1.1.3. Linguistic features

Ega shows an unusual linguistic complexity (Gibbon 1999, 2) and some of its interesting and unusual linguistic properties are outlined here.

Compared to other Kwa languages, Ega has more complex phonetics, phonology and morphology and also shows syntactic characteristics that are not typical of Kwa. (Connell/Ahoua/Gibbon 2002, 99)

1.1.3.1. Phonology

The Ega phonemic inventory consists of consonants and vowels. It includes the following consonants (Connell/Ahoua/ Gibbon 2002, 100):

Table 1: Ega consonant inventory

	Bilabial		Alveolar		Palatal		Velar		Labial Velar	
Plosive	p	b	t	d	c	J	k	g	kp	gb
Implosive		b<		d<		J<		g<		gb<
Nasal		m		n		/J		N		
Fricative	f	v	s	z			x			
Approximant						j				w
Lateral Approximant				l						

What is remarkable about the Ega consonant system is the extensive set of implosive consonants at five places of articulation. (Connell/Ahoua/Gibbon 2002, pp. 100) It is one of the most complete and still active series of contrastive voiced implosive consonants. (Ahoua/Connell/Gibbon 2000, 2)

The voiced obstruents in Ega function as depressor consonants, that is they effect tone realisation in Ega as explained below. (Connell/Ahoua/Gibbon 2002, pp. 100)

The following vowels are included in the Ega phonemic inventory (Connell/Ahoua/Gibbon 2002, 101):

Table 2: Ega vowel inventory

<table>
<tr><th>+ATR</th><th>–ATR</th></tr>
<tr><td>I</td><td>I</td></tr>
<tr><td>u</td><td>U</td></tr>
<tr><td>e</td><td>E</td></tr>
<tr><td>o</td><td>O</td></tr>
<tr><td colspan="2">a</td></tr>
</table>

The system of the nine vowels in Ega is divided into two sets according to tongue root advancement (ATR). This is the basis for vowel harmony in Ega, which means that there are consistent ATR values within a prosodic word. (Connell/Ahoua/Gibbon 2002, pp. 101)

Additionally, there is complex vowel hiatus (Ahoua/Connell/Gibbon 2000, 2), i.e. the occurrence of two consecutive vowels forming separate syllables. Vowel hiatus is a very frequent phenomenon in Ega and often results in phonological processes such as elision, coalescence or assimilation of one of the two vowels. (Connell/Ahoua/Gibbon 2002, 100). This affects the tones of the words and leads to contour tone although Ega uses mainly register tones.

Ega is a register tone language and uses the three basic level tones High (H), Mid (M) and Low (L). Tone functions both lexically (examples b) and c)) and grammatically (examples a) and b)):

a) nI´ ta´ I speak

b) nI´ ta- I spoke

c) nI´ ta` I chew

Tone realisation is relatively uninfluenced by declination. Still, tonal contours arise through two different processes (Connell/Ahoua/Gibbon 2002, pp. 102):

- combination of two tones on a single syllable after vowel deletion leading to HM, HL and MH contours
- after depressor consonants

1.1.3.2. Morphosyntax

Ega has the most complete nominal and gender class prefix system among the Kwa languages. Nouns and nominal phrases are discussed in more detail below.

Additionally, Ega shows complex verbal units with an intricate tense/aspect system which is also rarely documented among Kwa languages. (Ahoua/Connell/Gibbon 2000, 2) Aspects which play a role in Ega verbal units are the verb stem, tone, affixes, vowel apophony, an utterance final negation particle and subject and object pronouns. Table 3 gives examples for some of the relevant aspects in Ega verbal constructions.

Table 3: Examples for Ega verbal morphology

Aspect	Verbal Construction	Example
Tone	Tense/aspect/ accomplishment	nI´ pE` *I buy* vs. nI´pE- *I have bought*
Apophony	Conjugation	cI` *to laugh* vs. ni´ cO` *I laugh*
Prefix	Tense: future prefix a-	gi-ge´ pE`sa´ka´ *Guigue buys rice* vs. gi-ge´ a- pE`sa´ka´ *Guigue will buy rice*
Suffix	Tense: past suffix li´or lI´	nI´ pE`sa´ka´ *I buy rice* vs. nO´- pE- lI´ sa´ka´ *I have bought rice*
Negation particle	Negation	nI´ pE`sa´ka´ *I buy rice* vs. nI´pE- sa´ka´ a´ *I don't buy rice*

1.1.3.3. Conclusion

Given the unusual linguistic features of Ega, Gibbon (1999, pp. 8) draws the conclusion that Ega appears to be by far the most conservative of the Western Kwa languages. To some extent, the unusual linguistic features Ega has may, as mentioned above, reflect more archaic stages of Kwa language development. (Connell/Ahoua/Gibbon 2002, 99)

1.2. The Ega people and their sociolinguistic situation

1.2.1. The 1999 sociolinguistic survey

To get a more precise view of the sociolinguistic situation of Ega in the villages, Connell/Ahoua/Gibbon (in press) carried out a survey consisting of a series of questionnaires in some of the Ega villages in 1999. The survey investigated language knowledge and use among the Ega and followed a methodology used in similar work on other African languages. It aimed at determining to what extent the language is still actually used, which functions it serves and whether it is being transmitted to younger generations. Basically, the idea was to find out who speaks what to whom and when. (Connell/Ahoua/Gibbon, in press)

The first questionnaire is situated at the village level and provides a profile of the village in question. It is conducted with the chief or another knowledgeable person. It looks into which ethnic groups there are in the village, the degree to which Ega is spoken, the other language(s) spoken in the village, the language(s) spoken at school

and the lingua franca used in the village. (Connell/Ahoua/Gibbon, in press)
The second questionnaire is administered as a door-to-door survey and results in a linguistic census of the village. This questionnaire includes

- factual variables, mainly meta data about the informants
- attitudinal variables, which include the usage of the various languages in various situations/scenarios
- a "wish list" asking whether the speakers want their respective language to be used at school and whether they want their language to be put in writing

(Ahoua/Connell/Gibbon 2002)
The third questionnaire investigates the knowledge children have of their language and in which situations they use their own language or another language in its place. It is conducted with school-age children on a one-to-one basis. (Connell/Ahoua/Gibbon, in press)
The language questionnaire was administered in selected villages by two Ega speakers (male, mid 20s) trained in questionnaire techniques. (Ahoua/Connell/Gibbon 2002)
The results of the questionnaire with respect to Ega endangerment will be discussed in more detail below. Still, a short summary is given here in order to lay the basis for some comments and criticism relating to the survey.
In short, the general conclusion Connell/Ahoua/Gibbon (2002) draw is that Ega certainly is an endangered language. Their working hypothesis on Ega language endangerment of Ega is that language shift among the Ega is age-dependent, peripherality dependent and register-driven.
Ega endangerment is age-dependent because the intergenerational transmission is low in many, if not in all Ega villages. In the peripheral geographical areas there is a greater decline in than in the central areas. The language also serves a decreasing number of sociolinguistic functions, to the extent that in at least some villages Dida has replaced Ega as the primary language of daily use. Gibbon/Ahoua/Connell (2002) stress that this process is intensified as the mothers speak a dominant language because of exogamous marriages traditions of Ega men, invariably with Dida women. Finally, there is the spread of a dominant trade language (Dioula), a dominant immigrant language (Baoulé) and there is a dominant administrative language associated with monolingual (French) language policy.

1.2.2. Shortcomings of the sociolinguistic survey

There are some comments I would like to make regarding the survey. The comments are concerned with

- basic categories
- the set-up, i.e. the choice of subjects, the research interest and whether the choice of subjects makes it possible to answer the research question
- the interpretation of the results
- the implications of the results for this thesis

The data which is available for this thesis consists of the door-to-door survey which was carried out in the Ega village of Gnama and of the Gnama children questionnaire. Therefore, most of the arguments and results discussed here relate to Gnama and its inhabitants.

1.2.2.1. Basic categories

When investigating language use among a particular group, it has to be defined who belongs to that group, who does not and why or on what basis this classification is made. Connell/Ahoua/Gibbon (in press) do not clearly specify their group of subjects. In a follow-up presentation (Gibbon/Ahoua/Connell 2002) this is not done either. As a consequence, the basic categories or working hypotheses of the investigation remain unclear. These may be obvious to the researchers involved in the work but they are less obvious to the outsider. It is these questions in particular which are essential background for interpreting the figures but are not answered in the description of the survey:

- Ethnicity – How can you tell that someone is Ega?
- Transmission of ethnicity – How does someone become Ega?
- Traditions with respect to marriages – Who marries whom in terms of ethnicity? What happens after the marriage, i.e. who moves to whose village?

As far as I know, answers to these questions have not yet been published. Still, it was possible to obtain answers from Gibbon in personal communication.

Regarding the first and the second question, there has not been any explicit research. The hypothesis is that ethnicity is closely related to the first language and that it is the father who transmits the language and thus the ethnicity.

Providing that it is indeed the case that language and ethnicity coincide, the data from the Gnama children survey can be used to investigate the transmission of ethnicity:

Table 4: Transmission of ethnicity

Ethnicity	Child's language	Father's language	Mother's language
Ega	78	77	50
Dida	0	1	23
Other	0	0	5

The numbers given are total numbers and only children who gave Ega as their language were included in the survey. There are no children saying that one of their parent's language was Ega but who did not give Ega as their own language.

Except for one child, the languages of all the children coincide with the father's language. The child who said that his/her father's language was Dida also gave Dida as the mother's language. As a result, it seems safe to conclude that it is indeed the fathers who pass on their language and, if language and ethnicity go hand in hand, also pass on their ethnicity. Hence, one can tell that someone is Ega if that person's father is Ega.

Regarding the third set of questions, Gibbon (personal communication) sees a preference for Ega men to marry Dida women. It is not a taboo to marry within the ethnic group. According to Gibbon, the folk explanation behind the exogamous marriages is that Ega men consider Dida women to be generally better and, in particular, stronger, healthier and better cooks than Ega women. An Ega man who marries a Dida woman is seen as a lucky person, someone who marries an Ega woman is considered unlucky. Compared to the strong conclusion that the mother speaks a dominant language because of exogamous marriage which was drawn in Ahoua/Connell/Gibbon (2002), these interpretations are considerably weaker and more modified. Similarly careful statements are made in Connell/Ahoua/Gibbon (in press), where the authors say that "it is reported that at least in some villages intermarriage between Ega and Dida is common". Nonetheless, it is the stronger conclusions which are discussed and criticized below as these are the conclusions which are publicly available.

Regarding the figures on how many intra- and interethnic marriages there are, I analysed the data from the Gnama village and the Gnama children questionnaire. I used the answers to the questions investigating the father's and the mother's language. The data was edited in order to

- include only couples with at least one Ega partner
- not include any couple more than once
- exclude couples where the data is incomplete

Table 5 shows the numbers for the various marriages. The first ethnicity is the husband's, the second is the wife's ethnicity.

Table 5: Ega marriages and ethnicity

	Ega/Ega	Ega/Dida	Ega/Other	Dida/Ega	Total
Absolute	189	40	7	1	237
Percent	79,7	16,9	3	0,4	100

In the children questionnaire, some of the men are listed with more than one woman. As the Ega are mainly Catholic (plus they still follow their animist traditions) and the Catholic Church does not allow polygamy, it is likely that this is because the women died. Still, this is only a hypothesis. The issue is not addressed by Connell/Ahoua/Gibbon (in press) and needs to be investigated further to be cleared up.

Either way, the statement that there is an exogamous marriage tradition for the Ega is not fully supported by the data from the Ega village of Gnama. The majority of marriages there are within the ethnic group. Only one fifth of the marriages are between ethnic groups and only just over one sixth is between Ega and Dida.

The diverging outcome may be due to the fact that the data available for this thesis covers only one village whereas Connell/Ahoua/Gibbon (in press) worked with data from five villages.

Regarding the last question, Gibbon (personal communication) states that it is the woman who leaves her house and moves into the husband's house and village. This view is supported by the analysis above. With respect to the interethnic marriages, it is almost always the husband who is Ega, i.e. likely to be from the village, and the wife who is from another ethnic group, i.e. likely to be from another village. There is only one case where the woman is Ega and the man is Dida.

1.2.2.2. Set-up and analysis

The title "Ega: a preliminary assessment of endangerment" and the header "Language knowledge and use among the Ega" (Connell/Ahoua/Gibbon in press) imply that it is the interest of the researchers to investigate the Ega ethnic group and the Ega linguistic community. However, the exact research question is not given in either Connell/Ahoua/Gibbon (in press) or in Ahoua/Connell/Gibbon (2002). Also, the structure of the subject group is not given. Taking the village of Gnama as an example, the village profile questionnaire found that the village's population consists of Ega, Gurunsi, Dioula, Baoulé, Guro and Mossi. Then, 258 respondents answered the language knowledge and use questionnaire but it is not said whether these respondents are Ega or from another ethnic group. As a consequence, it is difficult to tell whether the research question can actually be investigated with the set-up.

Overall, it seems that the survey was in fact conducted on a regional, geographical basis. The inhabitants of the villages which are considered Ega villages were interviewed, no matter whether they actually are Ega or not. However, Ega men and women who do not live in these villages but may have for example migrated were not interviewed and thus are not included in the survey. This impression is confirmed by Gibbon (personal communication).

In other words, the survey did not investigate the Ega ethnic group but the Ega geographical region with all the ethnic groups living there. Again, this is a fact which may be obvious to the researchers and even intended by them but it is not instantly clear for the reader and the title of the survey may thus be misleading. It would make things clearer if the researchers stated what they mean by *Ega* – the geographical region, the ethnic group or the language. After all, the term *Ega* is used polysemically for all three concepts.

The geographical focus inevitably means that only a subset of the Ega community has been included. Therefore, the survey is partial and its results can only be interpreted against this regional background. It is not advisable to make statements about the Ega ethnic group as a whole on the basis of this data. In the long run, it is necessary to redo the survey and to include more members of the community. In particular, it is important to investigate the situation of Ega women who married Dida men and now live in neighbouring Dida villages. Extra questions should be included specifically tailored to look into the situation these women are in and these women's possible interests relating to the language, e.g. whether they are interesting in keeping their language up or in learning to read and write in Ega.

1.2.2.3. Interpretation of the results

The interpretation of the results leaves the impression of a rather single-sided and a not very objective point of view. Ahoua/Connell/Gibbon (2002) come to the conclusion that "the mother speaks a dominant language because of exogamous marriage".

Apart from the fact that this is by no means the average situation in the families, at least not for the Ega village of Gnama (cf. above), I see another problem in this interpretation. If there were exogamous marriages among Ega men as implied in the conclusion, then there would have to be the same preference for Ega women with respect to Dida men. After all, if Ega men are not supposed to marry Ega women, how can an Ega woman marry an Ega man?

As a result, if the Ega woman marries someone who is not Ega (and therefore likely to be Dida) there will be the constellation that the father speaks the other, dominant language, and the mother speaks Ega. This obvious fact has been left out by Ahoua/Connell/Gibbon (2002).

Dida women who marry Ega men move to Ega villages and bring their dominant language into a community that uses Ega and where the father of the family speaks Ega, too. However, the Ega woman who marries a Dida man takes her language to the dominant environment. Therefore for her, the situation is totally different. The possible constellations for interethnic Ega/Dida marriages are summarised in Table 6:

Table 6: Language contexts for Ega/Dida families

Husband/wife	Mother's language	Father's language	Village language
Ega/Dida	Dida	Ega	Ega
Dida/Ega	Ega	Dida	Dida

The chances for Ega to be used and to survive are likely to be higher in the Ega village than in the Dida village. The language situation of the women in non-Ega villages has been neither recognised nor investigated, although I imagine it must be very different from the situation in the Ega village.

Taken together, it is not justified to generally claim that the mother speaks the dominant language. Furthermore, it is only half the truth to state that there is an exogamous marriage preference for Ega men.

1.2.2.4. Consequences

The consequences that I see for this work is that first of all not all Ega have been interviewed. Not all the villages are included in the survey and, in addition, an important part of the Ega ethnic group, namely the Ega women married to Dida men and thus living in non-Ega villages has not been interviewed either.

Secondly, there are groups included in the survey that are less relevant for a thesis investigating the requirements for a primer for Ega as a native language. Therefore, if the survey is to be used as a basis for a definition of learner groups, it has to be read with a pinch of salt.

The results of the questionnaire with respect to context of language usage within given villages are still useful. They show the usage and decline of usage of Ega, which will be needed as a basis for the content of the primer, and they do not depend as much on the various ethnic groups and on whether all Ega have been included or not.

1.2.3. Contexts of Ega language usage

The Ega village of Gnama is chosen here as an example for investigating the contexts in which the various languages are used in Ega villages. Gnama has been chosen for the fact that this is the only data available for this thesis. The sociolinguistic data gathered in the village is included in the appendix, names have been edited out.

Figure 2 gives an overview of the major languages spoken in the Ega village of Gnama and the context of their usage for adult speakers. In cases where the sum of the percentages for one question exceeds 100 this is due to multiple answers.

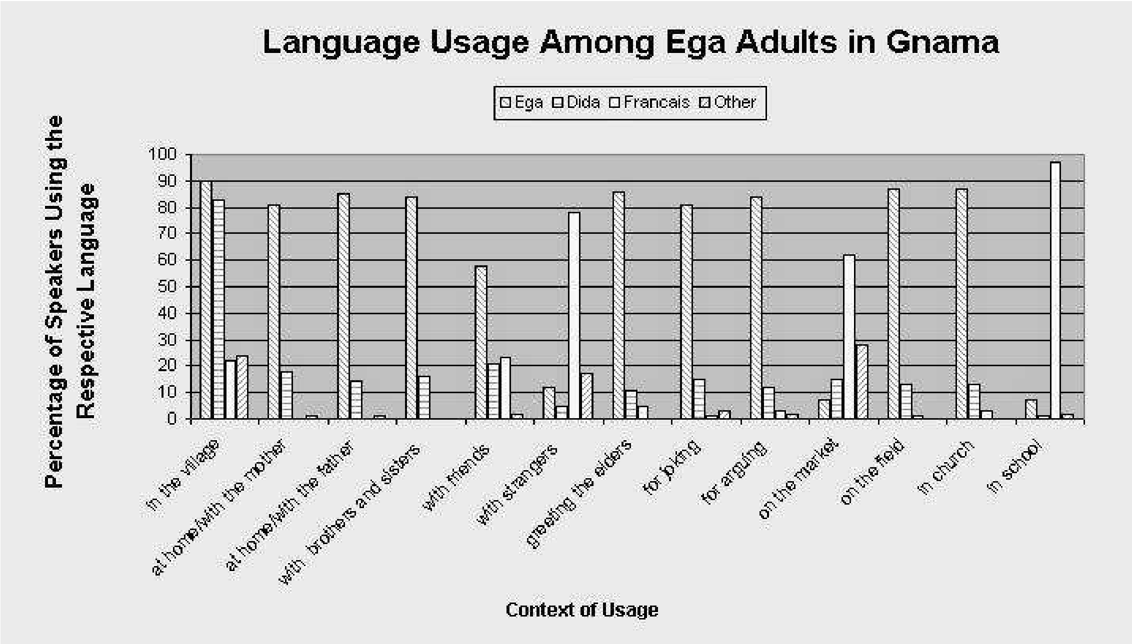


Figure 2: Language usage among Ega adults in Gnama

Figure 3 gives the same overview, this time, however, for children. Again, in cases where the sum of the percentages for one question exceeds 100 this is due to multiple answers.

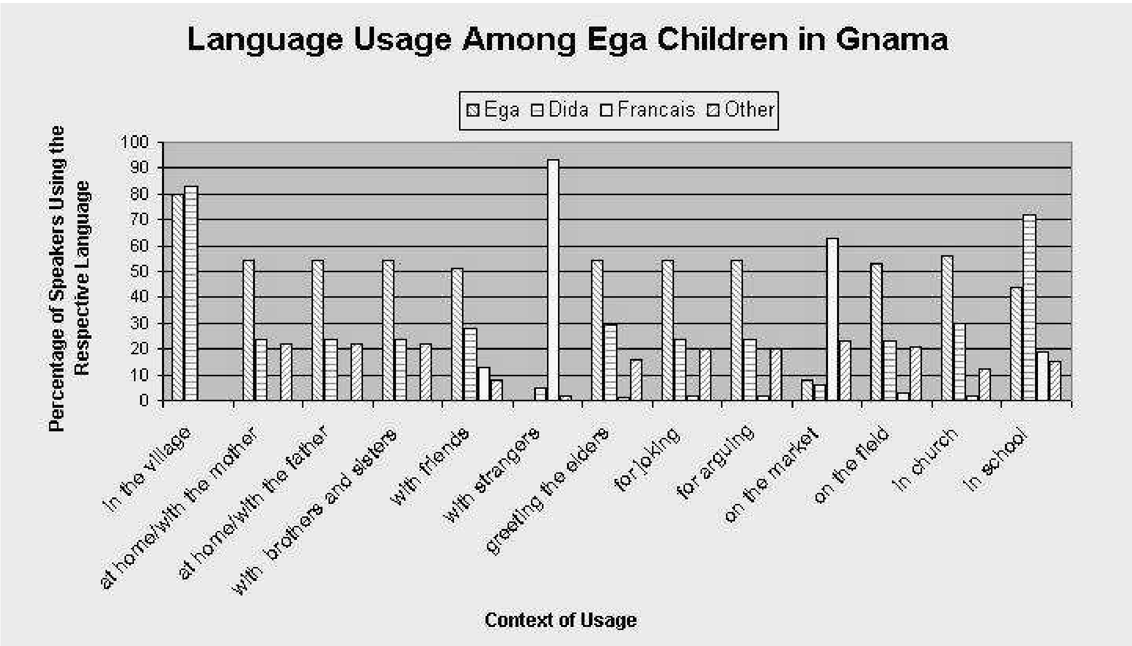


Figure 3: Language usage among Ega children in Gnama

Among adults, 90% name Ega as one of the languages spoken in Gnama, 83% name Dida. French is also mentioned and so are a number of other languages, namely Baoulé, Dioula, Goudou and Gouro.

Contexts where Ega is used by more than 80% of the village's adult population are

- at home with the mother and the father
- with brothers and sisters
- for greeting the elders
- for joking
- for arguing
- on the field
- in church

A context where just over half the village's adult population use Ega is when they speak with friends.

Contexts where Ega is used by around 20% and less of the village's adult population are with strangers, on the market and in school.

For the children, the situation is slightly different. 80% of them name Ega as one of the languages spoken in the village, whereas 83% name Dida, which means that from the children's point of view Ega has lost its status as the most widely spoken language in the village. Dida may exceed Ega by only 3% but the dynamics can be observed clearly, especially when the figures are compared to the figures for the adults, which are 90% Ega vs. 83% Dida.

The contexts where Ega is spoken and the percentage of children using Ega in these contexts are different, too.

There is no context where Ega is used by more than 80% of the children.

Contexts where Ega is used by more than 50% of the children are

- at home with the mother and the father
- with brothers and sisters
- with friends
- for greeting the elders
- for joking
- for arguing
- on the field
- in church

In school, Ega is used by just over 40% of the children.

Contexts where children rarely use Ega are with strangers, where Ega is used by 5% of the children, and on the market, where 8% of the children use Ega.

Table 7 summarises the contexts of usage among adults in comparison with the context of usage among children.

Table 7: Contexts of Ega usage among children and adults

Percentage of Ega used	Adults	Children
> 80%+	at home with the parents with brothers and sisters for greeting the elders for joking for arguing on the field in church	
~ 50%	with friends	at home with the parents with brothers and sisters with friends for greeting the elders for joking for arguing on the field in church in school
< 20%	with strangers on the market in school	with strangers on the market

The table shows that there is a shift towards less usage of the language and thus loss of the language in almost all contexts where Ega is still used widely by adults but not that widely by children. The exceptional context is with friends.

In addition to that, there are more children who use Ega in a school context than there are adults who think that Ega is used in school. There are two explanations for this deviation. Firstly, the adults may never have been to school and, unless they work there, do not go to school anymore. Therefore they may simply not know which languages are spoken there. The answers they give may be based on what they assume to be the language spoken there. Secondly, it may be that their answer is based on what they experienced when they went to school. Before, the monolingual practice in Ivorian schools may have been stricter. For example for the Ivorian language Agni, mother tongue teaching has only begun over the last years, which implies that indigenous language teaching did not exist before. Either way, it is good to know that at least in this context the use of Ega is apparently not decreasing.

1.3. Ega endangerment

The sections above gave an introduction to Ega as a language and to its sociolinguistic situation. The questions in this section are:

- Is Ega an endangered language?
- If yes, why is Ega an endangered language?

To answer the first question reliably, the term *endangered language* is defined more closely and then applied to the sociolinguistic situation of the Ega language. Following that, the factors contributing to the endangerment of the Ega language are discussed.

1.3.1. Endangered languages

Ega has been characterised above as an endangered minority language. There are a number of technical terms which are used by scientists when talking about languages like this, e.g. *linguistic minority*, *minority language* and various types of *endangered language*s.

LINGUISTIC MINORITY refers to the relative part which speakers of a language constitute within the total population of a country, e.g. the Francophone Canadians in Canada. These Francophones are considered a linguistic minority despite the fact that French is a language with many speakers and that it has the status of an official language in many countries. (Coulmas 1984, 10)

MINORITY LANGUAGES are defined as “minor languages that do not serve as a standard or national language in any country” and, more specifically, “languages that, in the nation where they are spoken, are not an appropriate means of vertical mobility and full participation in national life”. (Coulmas 1984, 10)

When speaking about minority languages, the number of speakers is regularly considered not a sufficient and functioning criterion. The number may be arbitrary and give a wrong impression of the relative part that the speakers of a language constitute within the total population of a country. (Bamgbose 1984, 21). Srivastava (1984) criticizes this criterion for the same reason. Srivastava (1984, 100) adds power as a criterion for finding minority languages. Srivastava uses English in India as an example – although the number of native speakers is small, English is not considered a minority language because of the power and influence of the speakers of English.

To sum up, minority languages are characterised by one or more of these features:

- little power and influence in the society/nation

- low literary, political, national or educational status
- not a widely understood language
- small number of speakers

Regarding the endangerment status, the UNESCO Red Book of Endangered Languages (Salminen 1999) distinguishes between six categories:

- EXTINCT languages: other than ancient ones
- NEARLY EXTINCT languages: maximally tens of speakers, all elderly
- SERIOUSLY ENDANGERED languages: a more substantial number of speakers but practically no children among them
- ENDANGERED languages: some children speakers at least in part of their range but decreasingly so
- POTENTIALLY ENDANGERED languages: a large number of children speakers but without an official or prestigious status
- NOT ENDANGERED languages: safe transmission of the language to new generations

Ega is in a minority position with respect to the relative part that the speakers of Ega constitute within the total population of the Côte d'Ivoire which is 15.7 million (Mabe 2001, 121). This makes the Ega speakers a linguistic minority.

With respect to its status, its number of speakers, its scope of usage and the power and influence of its speakers it does not range high. Therefore Ega is certainly a minority language, too.

Finally, Ega is an endangered language in some of the Ega villages. In Gnama for example, there are still children speakers but their number is decreasing. In other villages, Ega is in a seriously endangered position, for example in villages at the periphery of the Ega region because there there are no children speakers left. (cf. Connell/Ahoua/Gibbon 2001)

1.3.2. Factors in Ega endangerment

Ahoua/Connell/Gibbon's (2002) working hypothesis on language endangerment of Ega is that language shift among the Ega is age-dependent, peripherality dependent and register-driven. I do not doubt that these factors are relevant but I will still discuss Ega endangerment in terms of different factors because in my opinion the factors above are not precise enough. I agree that register plays a role but in addition I will discuss the

factors prestige, transmission and traditions and I will combine age and region into the factor number of speakers. While I am aware of the fact that all the factors contributing to Ega endangerment are interrelated, I will still examine them separately.
An additional problem regarding the factors is that it is difficult to determine whether they are causes, or symptoms. Let us take the number of speakers or functions the language still serves as an example. The language decreases with respect to the number of speakers and its functions. Therefore less people are interested in speaking it so that the language becomes even more endangered and is even less used. If it is less useful, less people will speak the language and so on. Ahoua/Connell/Gibbon (2002) do not address this problem and do not say whether for them the factors are causes or effects. However, this question is important for planning language maintenance work because it is more effective to work on the causes than on the symptoms.

1.3.2.1. Number of speakers

With respect to number of speakers, a distinction has to be made between absolute and relative numbers. Absolute numbers may be significant but cannot be the only criterion for endangerment. Large communities may be endangered, small communities like the Basque speaking community may be relatively stable. Therefore, relative numbers are equally necessary. Examples for what can be expressed in relative numbers are

- the number of speakers in different age groups
- the geographical distribution of the speakers
- the ratio of speakers vs. non-speakers

(Ahoua/Connell/Gibbon 2001; Coulmas 1984)
In addition to quantity, the quality of the command the speakers still have over their language is an important criterion. Broadly speaking, there is a distinction between speakers, semi-speakers and non-speakers. Furthermore, researchers (e.g. Aikhenvald 2003) distinguish between perceptive and productive skills, which adds speaker-hearers vs. hearers only to the spectrum.

There is a tendency that the command of Ega differs among the various age groups. Still, this does not hold for all the villages. In some of the villages only the old people speak the language, in others all age groups do. (Connell/Ahoua/Gibbon 2001)
Migration also plays an important role. There is both migration within the ethnic group,

i.e. Ega migrating to other areas, and migration of other ethnic groups, i.e. immigration of other ethnic groups into the Ega area. Ega emigration concerns part of the male Ega population, who migrate in search of work. In addition, Gibbon (2001) found that part of the population may migrate due to climate or environment changes such as drought, reservoir, and deforestation. Immigrants from other ethnic groups who migrate into the Ega territory bring their languages with them so that the villages become more and more ethnically mixed. Eventually, this means that the ratio of speakers vs. non-speakers changes in direction of the non-speakers. Geographically speaking, this is the more so the more the Ega village is situated at the periphery of the Ega territory.

To sum up, Ega is endangered not only in terms of absolute numbers but also regarding relative numbers. Endangerment in relative numbers of speakers is mirrored in

- a drop of language usage in the various age groups
- migration which leads to a less homogenous geographical distribution of the speakers and, as a consequence, in
- a change in the ratio of speakers vs. non-speakers.

1.3.2.2. Functions of the language

The question of when people use Ega has been discussed in detail above. Therefore, only a few additional comments are made here regarding languages other than Ega and their contexts of usage.

Ega is often used in the family, especially in the villages with a majority of Ega as the population. Still, it is not the only language spoken in the families. The Ega live in close contact with their Dida neighbours and in a number of families one of the parents is Dida. Consequently, the Ega are perfectly fluent in Dida and often use Dida (and also French) as a public language. (Gibbon 1999, 11)

In school and for administration, the language employed by the government is French, which is also the only official national language. In addition to the usage of French in education and official communication, there has recently been the introduction of major Ivorian languages in schools, e.g. Baoulé or Bété. This effectively requires Ega speakers to be at least trilingual or quadrilingual. (Ahoua/Connell/Gibbon 2000, 2)

Last but not least, Dioula, the dominant trade language, is spreading and dominant immigrant languages like Baoulé arrive together with their speakers who migrate to the Ega territory.

1.3.2.3. Prestige of the language

The Ega perceive the Dida language as a stronger, dominant, more efficient language than Ega and Dida has higher prestige among the Ega. Accordingly, the perception is that the Dida community also dominates socially and politically. According to Ahoua/Connell/Gibbon (2000) this perception leads to submissive behaviour of the Ega with respect to the enclaving Dida community. The use of Ega is deprecated among Ega speakers themselves and some Ega characterise themselves as Dida speakers to outsiders. Apart from Dida, there are also other languages with a higher prestige than Ega, for example attractive migrant languages. (Gibbon 2001)

1.3.2.4. Traditions of the Ega as a people

The traditional ethnic history relates the Ega to the westward Kwa migrations through other Kwa areas (e.g. Abbey) as well as Kru areas. Thus, the ethnic history emphasizes the status of the Ega community as guests in foreign territories. (Ahoua/Connell/Gibbon 2000, 2) Whereas this is interesting for the linguist in terms of genetic affiliation of the language, it does not boost the image of the language or strengthen its position to be a guest and to not really belong there.

Also, there goes the supposition of a centuries old non-aggression tolerance pact with the Dida community (Ahoua/Connell/Gibbon 2000, 2), which ties the Ega even closer to the already dominating Dida.

Finally, there is the Ega preference of exogamous marriages among the Ega. Ahoua/Connell/Gibbon (2000, 2) state that these exogamous marriages are invariably with Dida partners, but the Gnama survey shows that other partners, e.g. Baoulé or Yacouba wives, are also possible. Either way, if there are interethnic marriages this implies that the families are at least bilingual, possibly with one parent speaking an already dominating language.

1.3.2.5. Intergenerational Transmission

Another factor contributing to the endangerment is whether the language is still transmitted to and learned by the children, i.e. the productivity of intergenerational transmission. (Ahoua/Connell/Gibbon 2002)

Ahoua/Connell/Gibbon (2001, pp.7) found that this question cannot be answered the same way for all the villages and depends a lot on the geographic situation of the village and the ratio of speakers vs. non-speakers in the village. The more at the centre of the Ega region the village is situated and the higher the number of speakers of Ega is in comparison with the number of speakers of other languages, the higher the degree of intergenerational transmission, i.e. the more children still learn and speak Ega. For example in the Ega village of Gniguedougou, the language is relatively stable. Ega is learned as a first language by the children and is used as a primary language in all aspects of daily life. However, in the village of Broudougou Penda, which is situated at the periphery and has only a small percentage of Ega speakers, Ega is now apparently not transmitted to the young. Instead, the children learn and speak Dida. (Connell/Ahoua/Gibbon 2001)

1.3.3. A future for Ega?

Although the language is endangered, Gibbon (2001) still sees a chance for language maintenance, re-vitalisation and development by introducing alphabetisation and literarisation, publication, education and the introduction of new terminology in Ega. Therefore, language learning is taken up in the next section.

1.4. Motivational scenarios for learning Ega

The central questions in this section are:

- Are the Ega interested in (re)learning and studying their language?
- Why would the Ega want to (re)learn and study their language?
- What linguistic background is there to build on?

1.4.1. Are the Ega interested in (re)learning Ega?

The general impression from work with the Ega people and from fieldwork stays in the village is that there is an awareness of the fact that the language is endangered and that there is an interest and effort in keeping it alive.

In addition to that, the data from the sociolinguistic survey by Connell/Ahoua/Gibbon (in press) supports this impressionistic evidence. In the survey, the subjects were asked whether they wanted their language to be taught in school and whether they wanted to be able to read and write in their language. At least in the village of Gnama (cf. appendix), all the adults who were interviewed and answered the relevant question wanted Ega to be taught and also wanted to be able to read and write in Ega. For the children, the results are related to the language they speak at home with their family and are given in Table 8:

Table 8: Gnama children and their interest in reading and writing their language in relation to the language spoken at home

Language	At home	In school	For literacy
Ega	54%	79%	79%
Dida	24%	1%	1%
Other	22%	20%	20%

The other languages spoken are mainly migrant languages and languages from neighbouring areas. Children who speak these languages at home also want to use these languages, i.e. their own language, in school and for writing.

The children who speak Ega or Dida at home together make up 78%, which is almost exactly the number of children who want to use Ega in school and for reading and writing. Further analysis of the data confirmed this impression – the Ega and the Dida children are interested in using Ega in school and for writing.

1.4.2. What do the Ega want to learn Ega for?

This question was not asked in the survey, therefore the answer cannot be given from an empirical or from an Ega point of view. From a language-teaching point of view, however, it makes sense to teach the language for the contexts that it is actually used in. As Ega is used as a spoken language only at the moment, a functional pragmatic approach focussing on the oral modality is suitable. This implies that the Ega are taught their language Ega so that they can use it orally in the contexts outlined in section 1.2.3.

As the Ega are also interested in using Ega in written contexts, they can (re)learn their language in order to read and write in it. Whereas writing does not pose that much of a problem as long as there is a pen and paper, reading does because at the moment there is nothing to read. Of course, as soon as one Ega starts writing there will be material for the others to read.

A third motivation for (re)learning the language can be inferred from impressionistic evidence gained in the Ega village: The Ega (re)learn their language in order to pass on their ethnolinguistic history and their traditions, e.g. the story telling.

1.4.3. Linguistic background

Linguistically speaking there are various concepts of “learning” an endangered language. The most frequently mentioned and important ones are

- language development
- language maintenance
- language revitalisation
- language reacquisition

These concepts are situated in a continuum between extinct languages and not endangered languages.

Language reacquisition takes place when the language was once known but has been forgotten and needs to be learned again. It is similar to language loss or attrition and reacquisition concepts in second language acquisition. Whether this is a relevant concept for the Ega cannot be said on the basis of the existing data, more research needs to be carried out to answer this question.

Language revitalisation refers to what Flores Farfán (2001, 188) calls an “almost terminal point at which some languages find themselves”, i.e. refers to almost extinct languages. This is the case in some of the Ega villages at the periphery of the region.

Language maintenance relates to languages which are also threatened by extinction but to a lesser degree than to languages for which revitalisation is needed. (Flores Farfán 2001, 188). This is the case for example in the Ega village Gnama, where Ega is still widely used by the adults but less so by the children.

Language development, according to Flores Farfán (2001, 188), suggests improvement of the language and movement forward. It involves for example coining new terms for new items like for example technical instruments. This is relevant for Ega parallel to language revitalisation and maintenance efforts. Speakers who are still competent can “update” their vocabulary and learners who start from scratch immediately learn vocabulary that can be used not only for traditional but also for today’s purposes.

This chapter gave an introduction to Ega, an endangered spoken minority language in Ivory Coast. The chapter outlined the sociolinguistic situation for the Ega language community, the language’s endangerment status and possible motivations for the language’s future development in terms of language maintenance and language teaching.

The next chapter takes up two issues mentioned here, namely language documentation and language teaching, sets them in relation to each other and discusses aspects relevant for both fields.

2. LANGUAGE DOCUMENTATION

2.1. Teaching and language documentation

There are three possible ways in which teaching initiatives and language documentation projects tie in with each other:

- documentation preceding teaching
- teaching triggering and thus preceding documentation
- teaching and documentation as parallel processes

The first overlap is hinted at in the definition of teaching given by Lehmann (2001, 5): Teaching is the

> presentation of linguistic material which proceeds from simple and everyday phrases to complex and peripheral constructions and which helps the learner in assimilating the material and in stepwise building a competence

This makes language documentation a prerequisite for formal and institutional teaching (Lehmann 2001, 6) because documentation produces this exactly this kind of linguistic material and prepares the ground for the linguistic description. Both are needed before formal instruction can start.

During teaching or while preparing teaching materials, researchers may notice gaps in their knowledge and gaps in the documentation. Also, the process of teaching itself may raise questions which have not been addressed in the documentation. Noticing gaps may trigger more documentary work and research. If this is a one-off occasion, it leads to a scenario in which teaching precedes documentation.

In contrast to this, there is a scenario of on-going parallel processes where impulses from the one activity are taken up by the other and vice versa. This circular process is illustrated in a model of language documentation and its logistical structure given by Gibbon/Connell/Ahoua/Urua (2001, 8):

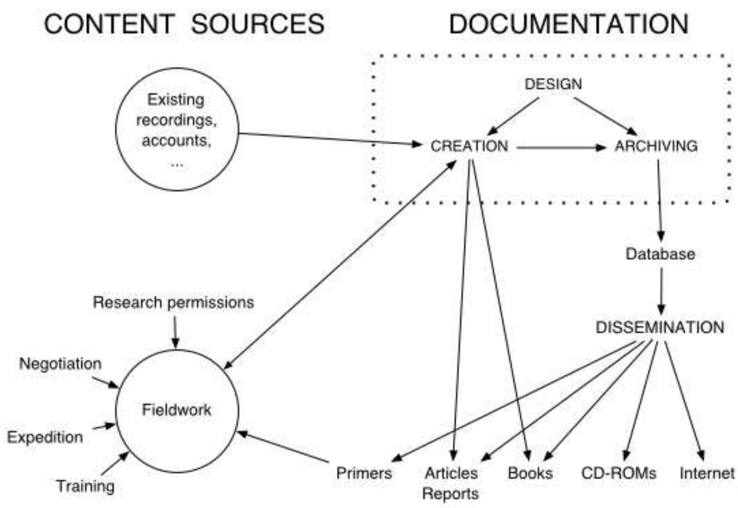


Figure 4: Language documentation and its logistical structure

In addition to perceiving teaching and language documentation as separate processes which mutually influence each other, language teaching can also be seen as a component and an outcome of documentation. In that case, it is a way of involving the respective community and is a method in language maintenance. This approach to language teaching and language documentation is expressed in the operational objectives of the Ega language documentation project (Gibbon/Connell/Ahoua/Urua 2001, 6):

> the development of constructive working relations with the speaker communities, in particular by provision of orthographic primers and reading material in the languages in cooperation with the village administration, and by training parents in the community in their use

Teaching endangered languages differs from for example teaching English as a foreign language in Germany. There are different constraints on the teaching situation of an endangered language like Ega because of

- the endangerment situation of the language
- the spoken character of the language
- the minority character of the language
- the infrastructural situation in developing countries
- the extent of the documentation and description

Language documentation can contribute to answering these questions and can give background information which is needed to see the whole picture.

In the same way, detailed, thorough and advanced teaching can only take place if the language in question has been documented in detail and maybe also described well.

Therefore this chapter discusses issues in language documentation and language description which are also important for primer making.

2.2. Language documentation vs. description

The central methodological question addressed in the documentation of endangered languages is how language documentation should be carried out. Ahoua/Connell/Gibbon (2001) specify this question and ask:

> How can a documentation be created whose quantity, quality and granularity guarantee its usability by future generations of the community, of scientists and of educationalists after the death of the language, that is in the permanent absence of native speakers and consequently without the classical interpretative and introspective methods?

The starting point of language documentation is descriptive linguistics; Himmelmann (1998, 161) suggests the application of the framework of descriptive linguistics to the recording of little known or previously unrecorded languages. In addition to descriptive linguistics, Himmelmann (1998, 167) lists a number of related disciplines which also have the potential to contribute to language documentation:

- theoretical and comparative linguistics
- sociological and anthropological approaches to language
- discourse analysis, spoken language research and rhetoric
- language acquisition
- phonetics
- ethics, language rights and language planning
- field methods
- oral literature and oral history
- corpus linguistics
- educational linguistics

It goes without saying that the list of contributors does not have to stop there. If documentary work is carried out on a larger scale than purely linguistic documentation, it is possible and productive to work together with musicologist, biologists, historians, archaeologists etc. (cf. the DFG Project "Kulturentwicklung und Sprachgeschichte im Naturraum Westafrikanische Savanne"[3]). Once this has been successfully done, why not take it one step further, leave the realm of the university and invite craftsmen and women to join in? This approach creates a very comprehensive model of language documentation which works truly interdisciplinary and combines language, culture, environment and any other aspect considered relevant for the language to be documented. In the end, it may even lead to a primer which includes vocational training making traditional artefacts such as traditional Ega baskets and chairs.

Taken together, the potential developments clearly show that language documentation is more than descriptive linguistics. This view is shared by Himmelmann (1998, 190), who calls for language documentation and documentary linguistics to be conceived as a field of linguistic inquiry and research in its own right, independent of for example descriptive linguistics. According to Himmelmann (1998, 167), the challenging subjects in this new field are

- to investigate and systematise the mutual dependencies within the field
- to find out about the relative importance of the various fields and
- to create a coherent framework for language documentation from these fields.

2.2.1. Defining the terms

Lehmann (2001, 7) defines language description as

> an activity (and, derivatively, its result) that formulates, in the most general way possible, the patterns underlying linguistic data

Language description is concerned with meta-linguistic knowledge. It provides interpretations and systematisations of the language as a system of abstract elements, constructions and rules which constitute the invariant underlying structure of the utterances observable in a speech community. (Himmelmann 1998, 164) The aim of language description is to enable the user of the description to understand how the language works. (Lehmann 2001, 7) The linguistic field which investigates language description is called descriptive linguistics.

Language documentation, on the other hand, is defined by Lehmann (2001, 1) as

> an activity (and, derivatively, its result) that gathers, processes and exhibits a sample of data of the language that is representative of its linguistic structure and gives a fair impression of how and for what purposes the language is used

In other words, language documentation is concerned with linguistic behaviour and the linguistic practices and traditions of a speech community. (Himmelmann 1998, 166) Its

3 URL: http://www.informatik.uni-frankfurt.de/~sfb268/

aim is to represent the language for those who do not have access to the language itself. (Lehmann 2001, 7) Himmelmann (1998, 161) introduces the term *documentary linguistics* for the field of linguistics which investigates language documentation.

2.2.2. Differences and overlaps

Having made the distinction between descriptive and documentary linguistics, differences and the similarities between the two disciplines are now discussed in more detail.

Table 9 summarises the differences between the two fields based on Lehmann (2001) and Himmelmann (1998).

Table 9: Differences between language documentation and language description

Aspect	Language Documentation	Language Description
Starting point	Data-driven	Theory-driven
Generality	Specific object	Generalisations
Consists of	Representations of its object	Discourse on the object
Level	Object level	Meta level
Role of primary data	Central concern	Means to an end, exemplification
Abstractness	Concrete	Abstract
Procedures	Participant observation, elicitation, recording; transcription and translation of primary data	Phonetic, phonological, morphosyntactic, and semantic analyses (instrumental measurings, distributional test, etc.)
Result	Corpus of utterances; notes on observations and comments by speaker and compiler on a particular form or construction	Descriptive statements, illustrated by one or two examples
Presentation format	Organised around the data/documents	Following a specific analytical framework
Methodological issues	Sampling, reliability, naturalness	Definition of terms and levels, justification (adequacy) of analysis
Target group	Laypeople	Specialists

This summary shows that there are substantial differences between the two disciplines. As a consequence, Himmelmann (1998, 163) calls for language description and language documentation to be kept apart and gives the following arguments to support this view:

- A data collection and its analysis are not just simply two different ways of presenting the same information, there are various possible interpretations and descriptions of the primary data. Separating description and corpus take these different views into account in a better way.
- Comprehensive linguistic descriptive analysis is not the only possible approach to primary data, other disciplines can work on the corpus, too. To enable access for various disciplines, the corpus must be kept apart from the analysis.
- There is the danger of a relative neglect of the data collection, perceiving it as subordinate to the seemingly more important description. Having the corpus separately and independently lets it gain status.
- There is the danger of the description determining the documentation, that is determining the kind of data considered to be relevant.

Given these arguments, language documentation and language description seem mutually independent at least theoretically speaking. It appears to be possible to infer the grammar of a language from a good documentation and a good description must be sufficient to generate data. Practically, however, this is difficult, if not impossible, to accomplish. (Lehmann 2001, 8) Furthermore, instead of a clear borderline, there is substantial overlap between the two disciplines. On the one hand, data transcription and translation require at least a preliminary analysis on various levels (Himmelmann 1998, 163) and therefore documentation inevitably contains descriptive elements. Similarly, a thorough description makes use of examples to illustrate its descriptions and includes documentary elements such as for example the genetic affiliation of a language. (Lehmann 2001, pp. 8) As a consequence, it is more realistic to see the relation between language documentation and language description as one of “bilateral mutual dependency”. (Himmelmann 1998, 165)

In addition to these arguments, a primer itself can be seen as an application and combination of language documentation and language description. The primer depends on an overlap of the fields because it requires both a thorough description as a backbone and an exhaustive collection of data to use as examples and exercises.

2.3. Objectives in language documentation

The primary purpose of language documentation is to represent a language for those who do not have direct access to the language itself. (Lehmann 2001, 5) The documentation must also aim at including additional material such as

- explanatory material, e.g. graphics, objects
- interpretative material which is not inherent in the primary data but generated by the author of the documentation
- ethnographic information

(Lehmann 2001, pp. 6)

This extension leads to the production and storage of a corpus of data which may for instance be used to teach a language. (Lehmann 2001, 5)

The aim of the theoretical field of documentary linguistics is to discuss and theorise about theoretical and practical considerations relating to language documentation. (Himmelmann 1998, 172)

2.4. Issues in documentary linguistics

The major issues which need to be discussed in documentary linguistics are

- Which communicative events should be chosen for the documentation and how can they be systematised?
- How can the quality of the data and the gathering procedures be guaranteed?
- How should the data be annotated?
- Which technical problems are there and how can they be solved?
- Which ethical and legal problems may occur and how can they be solved?
- Who can and will pay for all the documentary work waiting to be carried out?

So far not enough in-depth empirical and theoretical exploration of these issues has been carried out. All the same, there are suggestions and ideas. Not presenting and discussing them gives an incomplete picture so I decided to include them. However, as a consequence, this section will contain a lot of modal verbs such as *may*, *might*, *should*, *could* and so on. The reader is kindly asked to excuse them.

2.4.1. Parameters for the selection of communicative events

What are the parameters in determining the kind and number of communicative events to be included in language documentation? The ideal answer to this question lies in establishing a universally applicable grid of text types for language documentation. To do so, it is advisable to combine two approaches – an anthropological approach and a linguistic approach. (Himmelmann 1998, 177) The anthropological approach works on the basic assumption that communicative events are organised in a culture-specific way. The linguistic approach, on the other hand, looks for structural similarities in the various text types which can be applied universally[4]. (Himmelmann 1998, pp. 177) Combining the two approaches enhances the chances for not missing too many text types and for ensuring the comprehensiveness and representativeness of a corpus of communicative events.

The first universal parameter to be employed in the grid for the selection of communicative events is SPONTANEITY, i.e. the amount of time available for planning one's verbal behaviour. This parameter constitutes a continuum ranging from unplanned to planned. Examples from this continuum are exclamations, directives, conversations, monologues or ritual speech event. This parameter is of course not yet sufficient to produce a comprehensive corpus but it is of major importance for making the compiler aware of potential gaps in the corpus. (Himmelmann 1998, pp. 178)

The second parameter is MODALITY, i.e. the way the speakers can convey a message. Modality is a categorial rather than a continuous parameter and includes for example the categories speaking, signing, or writing. (Himmelmann 1998, pp.182)

The third important parameter is VARIETY AND REPETITION. This means that on the one hand it is necessary to record many different communicative events in order to cover the spectrum of possible communicative events in a given language. On the other hand, it is also necessary to have a considerable amount of repetition in the corpus in order to determine what is "regular" and what is ad hoc in a given example. (Himmelmann 1998, 183) Linguistically speaking, what is required is a variety of types and for every type a range of tokens.

In addition to these parameters, Lehmann (2001) suggests using the Jakobson model of communication as a quasi template and taking its components as universal parameters

---

4 It could be argued that what Himmelmann calls an anthropological approach really represents a functional and sociolinguistic approach whereas Himmelmann's linguistic approach is more of a structural and descriptive linguistics approach. This implies that instead of combining two disciplines there is a combination of two fields within the one discipline of linguistics. A thorough discussion of this issue goes beyond the scope and needs of this paper. Still, the summary of general principles for primer making given in section 5.7 builds on this combination of approaches, be it a combination of research paradigms, disciplines or fields.

for the selection of communicative events. This would add the parameters SPEECH ACT PARTICIPANTS, CONTEXT, COMMUNICATIVE TASK, TOPIC and CHANNEL to the list.
A careful selection and documentation of communicative events is of central concern for primer making because, as it will be outlined below, actual communicative situations and functional use of language build the basis for a primer.

2.4.2. Quality of the data and the gathering procedures

Criteria for evaluating language documentation are discussed in detail below. Still, one general point is important to make here already.
As in all social sciences, there is observer's paradox to watch out for. This paradox expresses that the object of the research is susceptible to change because of the ongoing research process. (Himmelmann 1998, 184) This is very important to keep in mind and researchers simply cannot be too aware of that. Fortunately, there are ways to handle this phenomenon. One possible way is to develop suitable means of gathering data which keep down the influence of the observer. Another way which must be used in addition to any of the approaches is to include meta data on each and every recording so that the degree of the influence coming from the observer can be traced back or inferred. (Himmelmann 1998, 184)
Good quality data and comprehensive meta data is important for a primer aiming at teaching a language in the actual context of its usage. Real-life data and information on how and when (not) to use linguistic constructions are absolutely essential.

2.4.3. Annotation of the primary data

Lehmann (2001, pp. 9) suggests the following minimum number of levels to represent primary data linguistically:

- linguistic raw data (video films, audio recordings)
- representation of this data at various linguistic levels (phonetic, phonological, morphological, syntactic, …)
- translations (interlinear morphemic gloss, free translation)

These levels can be extended and more levels can be added – ideally, there would be as many levels as possible and as needed. The urgency of language documentation, however, forces researchers not to go into too much detail for the moment. In addition, the documentation should be as free as possible from descriptive theory, therefore description and analysis should be restricted to a minimum.

When the levels are examined in detail, there are still open questions. For instance regarding the translation, the researcher has to answer the questions what a useful translation would be for the given text, which target language is most suitable and how an adequate translation can be achieved in the first place. For the gloss, the question remains how to segment the utterances for the interlinear glossing.

2.4.4. Technical questions

In collecting, annotating, archiving and presenting the data to the public there is the question of the appropriate technology. Researchers have to find out which media and which programmes are best to use and how to provide and, even more importantly, how to control access to the documentation. (Himmelmann 1998, 189)
Another point is how to enable the actual speech community to access the data when they might not have the required technical infrastructure. A possible way here is to set up regional heritage archives in co-operation with local and regional institutions (Gibbon/Connell/Ahoua/Urua 2001, 7). The centres could be provided with the necessary equipment and could serve as quasi mirror corpora and archive all the data. The centres could be a means of including the community in the documentation process and could allow the community both to get hold of the data and to keep an overview of what has been done. This is relevant with respect to ethical and legal questions because how can communities control access to their data if they have lost track of what has been done?
Regarding community access to the data, one could argue that although it is an honourable intention to share all the information it may not be useful for the community as they may not have the knowledge to fully use the data. Still, in return it can be argued that researchers do not understand everything they record straight away either and still collect it. Accordingly, the community should at least be given the chance. Providing that the community is interested and that they receive the necessary technical support they are bound to find a way of making use of the data for themselves.

2.4.5. Ethical and legal questions

> The interests and rights of the contributors and the speech community should take precedence over scientific interests. (Himmelmann 1998, 172)

This directive must be a leading principle in any language documentation. It implies that the speech community has the right to impose restrictions on the documentation process and that the researcher must obey these restrictions.

Himmelmann discusses restrictions in terms of internal and external relations. Internal secrecy controls and inhibits the information flow within a given society. It is mostly motivated by secrets and taboos and may affect the extent and the public availability of the documentation. In terms of external relations, the documentation may be restricted in order to protect the community, to prevent the exploitation, ridiculing or improper portrayal of its (linguistic) culture. (Himmelmann 1998, 173)

A question closely related to the ethical issues is the question of active involvement and participation of the language community in the documentation process. Active participation is a way of building confidence; it may help communities accept the documentation more readily and may help to solve conflicts. In addition, it allows the members of the community to see themselves as participants in the working process and enables them to contribute to and shape the ongoing research. If the community is also interested in language maintenance work such as primer development and actually using it to teach the language, this may in return get the researcher involved in this area of work, too. (Himmelmann 1998, pp.188)

2.4.6. Funding

Funding is a substantial problem in language documentation and for teaching initiatives. Language documentation projects are long-term projects where four or five-year funding, as is usually given for projects, is by no means sufficient. Himmelmann (1998, pp. 189) therefore suggests that a foundation is established which explicitly focuses on and sponsors the documentation and maintenance of (endangered) languages.

2.5. Formats used in documentary linguistics

2.5.1. Traditional formats vs. formats in language documentation

Traditional formats of publishing data from language documentation are very ambitious, time-consuming and require a lot of linguistic expertise. They take very long to prepare and remain inaccessible for a long time, which can prove to be crucial for endangered languages as they are about to die out sooner rather than later. Therefore the current ways of processing and distributing fieldwork data must be improved and made more productive and efficient. (Himmelmann 1993, 67)

A possible way to achieve this is to allow for initial language documentation formats to be not as systematic or comprehensive as the established formats. This has the following advantages (Himmelmann 1993, 68):

- It would allow for "fast" results (dictionaries, word lists, texts, …) which may be more in the interest of the speech community than scientific grammars.
- It could be instrumental in preserving (or even rebuilding) self-esteem and cultural identity in the speech community due to the simple fact of seeing the language in print next to a standard language of high prestige.
- It would allow for calling *respondents* what they are – contributors to a compilation of language materials.
- It is a convenient way to make additional data accessible which has been collected but was in the end not actually needed by a particular researcher
- It makes it easier for researchers from other fields to use the data as these formats are easier to access and understand by non-specialists.
- It includes more "raw" data and less thoroughly edited or "tidied up" data and therefore allows for better transparency, re-evaluation and verification

Surely, the call for new, less ambitious formats in language documentation must not be misunderstood as a call for lower scientific standards for established formats. Instead it is meant to improve the image of less sophisticated formats. Sophisticated analysis can be at a later point based on the written sources provided that recording has been done and made accessible when there were still speakers around. (Himmelmann 1993, pp. 66)
In terms of a primer, this means that it is not necessary to wait until a whole series of textbooks and supplementary material has been set up. Instead, small modules can be designed which start with one issue and can be used straight away. As time goes by, more and more modules can be added so that the primer is compiled step by step.

2.5.2. Components of a document in language documentation

A document format in language documentation should include information on a complete communicative event. (Himmelmann 1998, 168) It should be a format that emphasizes via its structure a holistic and situated view on linguistic behaviour, which implies that for example it includes the communicative setting in addition to the raw linguistic data.
In addition to the individual commentaries for the various recordings which include information on everything which happened during eliciting, recording, transcribing and translating, Himmelmann (1998, 170) suggests a general introductory commentary. This

general commentary includes information on the speech community, the language, the fieldwork, the methods used for gathering and processing the data, the contents and scope of the documentation.

For commentary that cannot be allocated to a particular communicative event, Himmelmann recommends a further component which he calls *analytic matters*. This part accommodates elicited data or discussions of language matters independent of a particular communicative event.

Taken together, this sets the following components:

- raw data
- linguistic annotation
- individual commentary
- general introductory commentary
- analytic matters

This is already quite comprehensive but it still does not include Himmelmann's initial demand that background information on the communicative event must be given. This may be included in the individual commentary but for archiving and access reasons it is better to keep linguistic commentary and meta data apart and to set up an additional category "meta data of the recording". There could be cross-references going out from this to the individual recordings, which would be a very convenient database to search for the user. (cf. Gibbon 2002)

2.5.3. Examples for possible formats

The new formats in documentary linguistics could consist of slightly edited versions of field notes, texts, elicited lists (e.g. morphological paradigms, numbers, measures, folk taxonomies) or observations. Slightly edited in this case means that the data has been for example edited for spelling or glossing, but, as said above, the data does not have to be as systematic or comprehensive as required in the established formats. (Himmelmann 1998, pp. 165)

2.6. Quality criteria

Himmelmann (1998, 164) severely criticizes that descriptive linguistics, unlike other disciplines, has not yet profoundly dealt with methodological issues relating to obtaining and presenting primary data.

A primer requires quality primary data both as a source for its descriptive elements and its texts and exercises. Low quality data leads to a low quality primer as it cannot represent and teach the language the way it is used. Because the primary data is that essential, quality criteria for primary data are reviewed here.

2.6.1. General criteria

Lehmann (2001, 6) rightly points out that in language documentation it is an illusion to hope for true objectivity. The process of documentation already constitutes a selection of what part of the reality will or will not be included. Therefore including interpretation, i.e. a certain amount of subjectivity, need not be regretted or seen as affecting objectivity.

Reliability is a criterion which both Himmelmann and Lehmann mention; both do not explore it in detail. Looking at other disciplines, however, reliability is understood there as whether the research does actually investigate or measure what it aims at investigating or does whether it maybe measures at something different. This could be due to unsuitable tools or also because of the observer effect mentioned above. Possible ways out of the latter have been discussed above.

Regarding the documentation of endangered languages, Gibbon/Connell/Ahoua/Urua (2001, 8) add the criterion of explicitness. They postulate a “Principle of Explicit Documentation” and state:

> Language documentation is explicit if it enables insightful study of a language and culture when the language is extinct and when native speakers are no longer available as sources of creative, operational, structural and interpretative knowledge about the language.

Apart from the general criteria, there are further quality criteria relating to the individual steps in the process of language documentation.

2.6.2. On collecting the raw data

Firstly, researchers need to aim at the representativeness of the documentation. Representativeness means that the documentation “adequately represents the purposes and ways the language is made **use** of in the speech community and the structural properties and possibilities of the linguistic **system.**” (Lehmann 2001, pp. 11, bold font his) Lehmann softens this demand by agreeing that full system representativeness of a documentation cannot be guaranteed as this presupposes linguistic analysis. Furthermore, there is no widely agreed universal model to use instead of a preceding analysis because linguistic systems are language-specific. It is possible, however, to reduce this problem by using the parameters for the selection of communicative events discussed above. A documentation can be considered representative if there is sufficient variation on each of the parameters.

Secondly, a documentation should be complete in the various dimensions of representativeness outlined above. As with endangered languages time is a crucial factor, Lehmann recommends starting with the core situations and proceeding to the peripheral situations at a later stage.

Thirdly, Himmelmann (1998, 168) discusses comprehensiveness in addition to representativeness and completeness. Unfortunately, he does not give a definition of his usage of the term and therefore it remains unclear what he means by that and how this term is distinguished from the other two.

The fourth criterion is quality as proposed by Lehmann (2001, pp. 10). Lehmann distinguishes between quality of content and quality of form. Communicative events with high quality of content are events considered important for characterising the culture, e.g. myths, prayers or instructions to the youth. Routine everyday conversations on the other hand are considered less important with respect to the culture and therefore of a lower quality of content.

Quality of form is associated with linguistic correctness and with what Lehmann calls aesthetic beauty of a text, which is related to the linguistic skill of its authors. Literary texts, for example, fall into this category. The criterion of aesthetic beauty is, however, a very subjective criterion and hard to define and therefore I doubt that it can be included just like that.

The quality of spoken texts can be defined as a function of textual cohesion, stylistic appropriateness, richness in system resources, correctness of constructions, accuracy of articulation and, last but not least, the technical quality of the recording. (Lehmann, 2001)

Lehmann points out that quality and representativeness may conflict, for example when recording everyday routine communication. Routine communication is highly representative, but Lehmann says that it has low cultural value and concludes that the choice whether to include it or not depends on the purpose of the documentation. In my opinion, however, it is not justifiable to reject everyday routine communication and to see it as having low cultural value. Quite on the contrary, it is well feasible to infer a lot about the communicative behaviour of a people and about their conception of politeness from everyday routine communication. House (1996) for instance has investigated and found this for German and English routine pragmatic idioms. From the linguistic forms, she deduced politeness parameters characteristic for German and English respectively and where the values for the parameters differ. This shows that everyday routine communication has not necessarily little cultural value. Instead, it may well reveal a lot about the culture.

Finally, the last criterion in data collection is naturalness. Naturalness refers to the degree to which speakers are linguistically self-aware. It constitutes a continuum ranging from completely unaware to paying full attention to linguistic form. Points in this continuum are for example natural, observed, staged communicative events or elicitation. As completely natural data is difficult to obtain (cf. observer paradox), Himmelmann (1998, pp.184) stresses that more research needs to be carried out investigating the staging of communicative events and the elicitation of linguistic data. Also, it is essential to further explore the limits, possibilities, comparability, efficiency, usefulness and the culture-specificness of the staging and elicitation techniques and methods.

2.6.3. On the annotation

For Lehmann (2001, 11) it is most important that data editing is restricted to producing further representations only, if possible by engaging the authors of the texts themselves. By no means must the original recording itself be changed.

What has to be kept in mind concerning the annotation is that there must be a clear definition of terms and levels used in the annotation. Also, the preliminary analysis and description of the data must be justified and adequate. (Himmelmann 1998, 162)

2.6.4. On storage

The data should be recorded on an enduring medium to keep it for posterity (Lehmann 2001, 6) and it should be accessible for all its potential users independent of their background or interests. (Lehmann 2001, 5)

Now that the concept of the primer has been placed into the framework of language documentation, its linguistic basis is defined and described in the following chapters.

3. LINGUISTIC DESCRIPTION OF A PRIMER FOR AN ENDANGERED LANGUAGE

Finding a definition for the term *primer* proved to be difficult. Gibbon/Connell/Ahoua/Urua (2001, pp. 6) include the primer in their overview over documentation linguistics and additionally use the term *orthographic primer*. In both usages, however, they do not give a definition. Also, I did not find a definition in the current literature on language teaching. Modern teaching research and teaching initiatives seem to work on a larger scale and aim at establishing elaborate language syllabuses and curriculum guidelines, e.g. the Council of Europe's (2001) framework which includes detailed information on language learning, teaching, and assessment.

In a use which the Cobuild English dictionary (1995, 1305) describes as old-fashioned, *primer* is understood as a book containing basic facts about a subject and is used by someone who is beginning to study that subject. The Oxford Concise Dictionary (Pearsall 1999, 1135) goes in the same direction but extends the scope by defining a primer as a book providing a basic introduction to a subject or used for teaching reading.

A primer in the usage in this paper is understood as a collection of teaching material which introduces learners to a language and its usage. A primer can be in the form of a book, but does not have to be. It is intended for learners who begin to study a language although their starting points can be situated on a scale of language proficiency, i.e. there may be learners who are already proficient but want to look behind the everyday usage of their language into a meta level or they may be learners who start from the beginning.

The linguistic description of the primer is based on a textlinguistic model of documents. This model includes 3+1 components, which are related to each other:

- document CONTENT, which linguistically speaking outlines the semantic properties of the document. In terms of the primer it specifies what the learners will learn, i.e. which phonological, orthographic, morphological, syntactic structures, which vocabulary item and which kinds of discourse.
- document STRUCTURE, which linguistically speaking outlines the syntactic properties of the document. In terms of the primer, structure relates to the order in which the content is taught, i.e. the progression.
- document LAYOUT or rendering, which linguistically speaking outlines the

pragmatic or surface properties of the document. In terms of the primer, layout states which media and modalities are used for the primer, i.e. how the content is presented and which material is used for the presentation.

- CONTRAINTS on the document, which affect content, structure and layout. There are constraints due to the objectives of the teaching, the learner groups, the situation of the language, the primer makers and users and due to external constraints such as the political situation in the area where the language is spoken.

Figure 5 displays the components of the document and shows the dependencies and links between the various components.

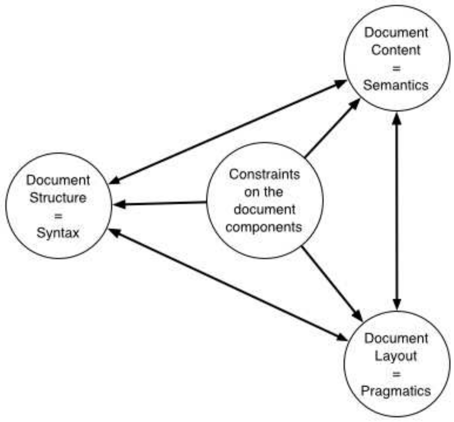


Figure 5: Textlinguistic model of the structure of a document

As the description is based on this model, the following chapters will follow the structure of this model, too.

Chapter 4 discusses constraints on the primer. It is followed by two chapters dedicated to the content of the primer. The first of these two content chapters discusses the content from a general linguistics point of view. It starts with the larger units in language, i.e. texts and discourse, which are usage in terms of the primer. Then it works its way down via semantics, i.e. vocabulary in terms of the primer, syntax and morphology. It ends with the small units in language, i.e. phonetics and phonology for the spoken modality and orthography for the written modality.

The second content chapter applies the general theoretical issues to the language Ega and describes Ega in examples. Of course, it is not possible to fully describe Ega as this is beyond the scope and the needs of this thesis.

Chapter 7 is concerned with the document structure, i.e. the overall progression in the primer and a possible structure within a unit. Finally, chapter 8 comments on the document layout, i.e. how to present the content in various media and modalities.

4. Constraints on primer making

4.1. Teaching objectives

The fact that the Ega DOBES project aims at developing teaching material has already been mentioned above. The question which has to be answered now is why teaching is necessary and what is to be achieved.

Abstractly speaking, the overall objective is language maintenance among (semi-) speakers and language revitalisation and reacquisition among interested non-speakers.

A prerequisite for successful language maintenance and revitalisation is increasing the prestige of Ega. This is necessary because the prestige of Ega as a language is low among the Ega ethnic group (cf. above). For other languages it has been reported that putting a language in writing and developing teaching materials had just this effect and the language gained in prestige. As a consequence, speakers and learners felt motivated and more encouraged to learn and speak the language. (Aikhenvald 2003, 130)

More specifically speaking, the objective of the maintenance and revitalisation work is that learners are able to communicate in their language again. To start with, it seems a sensible aim to enable the Ega to use their language in a number of well-defined every-day life situations as outlined in section 1.2.3 on context of language usage taken from the Gnama survey. To accomplish this, learners need to have access to enough vocabulary, grammatical structures, routines etc. to use Ega for their every-day lives.

The view that teaching contents must be based on topics and situations relevant for the people and their lives is shared by Kuhn (1986, 219). Kuhn works with South American Indians and most strongly emphasizes that Indians need Indian schools with native teachers and with their own curricula which reflect the issues the Indians are concerned with.

The second objective is to promote mother tongue literacy because the Ega are interested in reading and writing in their language. (cf. above)

The third objective relates to cultural studies, i.e. enabling the Ega to pass on their ethnolinguistic history and their traditions like for instance story telling.

For all three objectives, the main focus is on procedural knowledge, i.e. knowledge on how to perform a task or on how to solve a problem. The teaching can be considered successful when in the end learners can produce for example a proper sentence. For procedural knowledge it does not matter whether learners can also explain the grammatical rules underlying that sentence.

Nevertheless, there is also a need for declarative and meta knowledge on the language. Declarative knowledge is knowing basic facts and recalling stored information, meta knowledge is knowledge about knowledge, i.e. how knowledge and information is organised and how it can be accessed and used. Learners need declarative, meta and procedural knowledge in order to use resources on the language independently without a linguist or a teacher helping them.

As for teaching methods, the objective in teaching endangered languages is to integrate institutional and traditional learning and thus to involve the community in the language maintenance and revitalisation work. This goes in line both with the demand to combine linguistic and anthropological approaches and with the methodological demand to start teaching from what is there in the language community.

For a general outline of the objectives in teaching endangered languages these remarks are considered sufficient. Where necessary, the objectives will be specified in the following chapters.

4.2. Target groups

The potential group of learners interested in teaching materials for Ega is very heterogeneous. To name only a few, it is likely that there will be differences with respect to:

- command of Ega
- age
- gender
- other languages spoken
- Ega spoken as a surrounding language
- received formal education
- literacy in French

One way of dividing the potential set of learners into subgroups is having subsets of school children, adolescents, adults who do not speak Ega anymore but want to learn it, migrants, and future learners after the death of the language. This results in learner groups which are homogeneous in terms of age or origin, which is a usual practice in for instance German schools.

For an endangered language, however, this makes less sense. The main difference here is in the command of Ega and, consequently, in the learners' needs for language development, language maintenance, language reacquisition or language revitalisation.

The specific focus for this paper is on language maintenance in villages where Ega is still widely spoken by the adults but to a lesser extent by the children. This subset of learners has been chosen because there is reliable sociolinguistic data available on them. Moreover, the linguistic data that is accessible is not sufficient for absolute beginners and additional data cannot be gathered at the moment for reasons given in the section on external constraints on primer making.

4.3. Constraints due to the situation of the language

4.3.1. Prerequisites

So far Ega is a spoken language, it has not yet been put into writing. In order to design a written primer it still requires an orthography and materials for teaching literacy, i.e. an orthography primer. In addition to that, actual reading material is required to promote literacy – after all, how can reading skills be used and practiced if there is nothing to read?

4.3.2. Urgency

Ega is an endangered language. It has low prestige among its speakers and intergenerational transmission is low. Although these are not favourable circumstances, there is still interest in keeping the language alive. Therefore, maintenance work needs to start sooner rather than later, preferably when there are still competent speakers around to support the initiative. If revitalisation and reacquisition start only after language death, revitalising the language will be more difficult as there will not be a natural surrounding for actually using the language.

4.3.3. Official Support

Ega is a minority language. The community of speakers is small and has little political power. So far there has not been official political interest or effort in maintaining the language. Fortunately, however, there is interest in the region, there is for example a judge who is interested in setting up heritage centres for the languages and peoples.

4.3.4. Infrastructure

The Ega community lives in a developing country. Government support for the language is not existent and so far there is no money to pay special teachers for Ega. The school infrastructure is less developed than in Europe. There are no qualified teachers for Ega as a first, second or foreign language. It is not guaranteed that school is accessible to all learners as it may be too expensive, too far away, people may be needed to work and it is possible that not all parents want to send their children to school. In addition, it has not been investigated what the equipment is like in the school.

4.3.5. Level of Documentation

As languages constantly change, there will never be a complete documentation or description to base teaching material on. However, if the language has only been documented, described and analysed a little, then the teaching material, too, can only be preliminary. Structures not verified by the data or found in the documentation or description must not enter the primer as they may be wrong. Detailed and advanced materials require more detailed documentation and description.
A special case arises if the language is extinct before a sufficient amount of data has been collected. A similarly challenging scenario arises when it is impossible to record purpose-made material for teaching because of for example travelling restrictions due to a civil war. In those cases primer designers have to make do with whatever data is available.

4.4. Constraints due to the people involved

4.4.1. Primer makers

When a linguist and a native speaker work together on making a primer they both bring in their expertise and their gaps in their knowledge. It is likely that both expertise and gaps are in complementary distribution. Ehnert/Hornung (1986, 207) illustrate a similar scenario in a model describing the distribution of competence in foreign language teaching for specialised purposes. The version of the model given here has been adapted to the circumstances in developing teaching material for endangered languages.

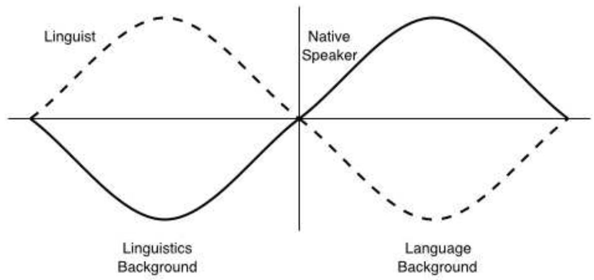


Figure 6: Distribution of competences in primer design

It is likely that the linguist has training in linguistics and maybe language teaching but lacks in-depth knowledge of and performance skills in the respective language. The native speaker has a command of the language but may not have a linguistic background. Therefore, they need to work together to achieve the best results.

4.4.2. Primer users

A primer has to serve the needs of its intended user groups. For more information on these groups see section 4.2 on target groups.

4.5. External constraints

Because of the unstable political situation in the Côte d'Ivoire it is currently not advisable to travel there at the moment. This means that it is impossible to develop purpose-made teaching materials and I have to make do with data which has been gathered for other purposes. Also, the teaching material cannot be double-checked for mistakes by native speakers and it cannot be tested and evaluated.

5. CONTENT OF THE PRIMER – LINGUISTICALLY SPEAKING

What to teach is a question which has been widely discussed in the field of applied linguistics and second language teaching. It is not, however, the purpose of this thesis to sketch this development. Instead, the focus will be on two factors influencing the choice of structures to be included in the primer – the teaching objectives and the methodological approach. A recent development in methodology for the context of European languages has been outlined by Council of Europe (2001) in its *Common European Framework of Reference for Languages*. Its action-oriented approach (Council of Europe 2001, 9) concentrates on usage and functions of the language and starting from that defines which skills a learner at a given level should have in terms of listening, reading, spoken interaction, spoken production and writing. Of course spoken endangered minority languages in a multilingual context are in a different situation than the major European languages focused on by the Council of Europe. Still, a pragmatic approach is equally suitable there given the teaching objectives defined in section 4.1. As a consequence, it is the functional use of the language which is at the centre of the primer content and the concept *functional use* can be operationalised as the who, where, when, what, why, how etc of language usage.

The context where functional use of language takes place is characterised by the domain, situation and topic. (Council of Europe 2001, pp. 45; Hopkins 2002, pp. 206) For so far undescribed languages, these domains, situations and topics can be identified by observation, by elicitation or, as done by Connell/Ahoua/Gibbon (in press), in a survey. Once typical examples of language usage in the identified contexts have been collected, documented and analysed, the rest of the content follows automatically. Vocabulary, grammatical structures, speech acts etc. which occur in the collected texts must also be part of the content of the primer.

The structure of this chapter reflects this view on the selection of primer content. It starts with usage, which is concerned with the larger linguistic units, and works its way down to phonology and orthography, which are concerned with the smaller linguistic units. Each section gives central concepts relevant for primer making, defines teaching objectives more specifically where necessary and discusses issues in the field. The reader will notice that the sections are not equally detailed. This is due to the fact that in the chapters which apply the linguistic issues to Ega and to the Ega primer only some aspects are dealt with. These issues require extra background, for the others only a rough outline is given.

5.1. Usage or Pragmatics

Usage of language describes how to do things with words. It includes everyday conversation just like it includes essential language functions such as greeting, making apologies, thanking, or asking for something. Some of these kinds of discourse are organised rather strictly, others are not. Some types of discourse such as greeting someone follow more or less the same underlying pattern in every society, but may look very different on the surface in the actual performance. (Eibl-Eibesfeld/Senft 1987, pp. 94) Therefore it is again necessary to approach usage from two directions – a systematic, model-based, top-down point of view while still keeping an open mind for observation and bearing in mind that there may be new forms which are not yet included in a given model.

5.1.1. Central concepts

Generally speaking, discourse refers to texts, i.e. a discourse is a linguistics unit consisting of several sentences. (Fromkin/Rodman 1998) For languages without a written modality it is the spoken discourse which is relevant. Spoken discourse in the form of dialogues takes place in a continuum reaching from freely composed to dialogues to ritualised communication[5] where most of what is happening is fixed.

For the free end of the continuum, linguistic documentation and analysis has to provide typical examples of communicative events spanning at least the relevant domains, situations, topics and speech acts. These exemplars form the basis for relevant linguistic structures for the content of the primer.

The ritualised end of the discourse continuum is characterised by a more rigid semantic frame, a more rigid semantic script and more fixed lexical material than the free end of the continuum. These three features situate the discourse in a given context, establish the order of the sequences and set the lexical material to be used.

A semantic frame is defined by Lüger (1999, 25) as a cognitive framework which takes a particular concept and relates it to what is typically associated with this concept by the language community. So basically, a semantic frame defines the where, when, who, why etc. of for example a typical dialogue for ordering a meal in a restaurant. Lüger (1999, 25) stresses that semantic frames are taken for granted by a language community – speakers assume that the frames are known by the interlocutors.

Whereas the semantic frame defines the external positioning of a piece of discourse, the semantic script describes its internal organisation. Lüger (1999, 25) defines the semantic script as a framework for interaction which describes standardised routines. Applied to the example of ordering a meal, the script defines whose turn it is to say something and what to say.

The lexical material employed in ritualised discourse is called various names: pragmatic idioms, routine formulae, gambits, fixed expressions, etc. The term used here is routine

[5] For a more thorough discussion of the concept of ritualised communication and its linguistic basis see Rafferty/Salffner (2002).

formulae. Routine formulae are defined by Koller (2003) as prefigured, fixed and established linguistic forms which are connected to

- certain communicative situations (e.g. greeting, thanking) or
- certain types of text (e.g. formal letters) or
- certain situations, actions, instructions (baptising, in court)

Routine formulae are characterised by the following features :

- They are highly standardised.
- They are highly predictable.
- They are ready-made templates and do not have to be synthesised for speaking, therefore there is highly automatised usage for them.
- They have a given function, which often is a phatic function.
- They have (partly) lost their lexical meaning. (e.g. in German *Mach's gut*.) (Coulmas 1981, Lüger 1999)

5.1.2. Objectives

The ability to communicate in free discourse in every-day situations has already been established above as an objective.

Ritualised discourse requires special attention. This attention is important because first of all both semantic frame and semantic script contain information on the cultural and ethnolinguistic heritage which may get lost if it is not entered in the curriculum. Secondly, the semantic frame is important to include because communication occurs in a context and by using a functional approach to language use as the basis of the teaching context is included by definition. For similar reasons, semantic scripts of discourse types have to be part of the curriculum. Finally, routine formulae are important to teach because the learners cannot compose them by themselves from smaller units. They have to be accessed and retrieved as a whole phrase.

To sum up, the objective in teaching language usage is achieving pragmatic performance and knowledge of language, i.e. knowing the right expression and being able to use it in the right situation and at the right time.

5.1.3. Issues in usage

5.1.3.1. Lexical material

This thesis separates routine formulae discussed in the usage section from other lexical items which are discussed in the vocabulary section. This raises the question why this has been done. After all, both routine formulae and other vocabulary are part of the lexicon of a language. They form a continuum reaching from prototypical free morphemes and stems at the one end to fixed phrases consisting of several words at the other end. So structurally speaking, they are not distinct subsets or categories within a larger set but mere extremes of a continuum of lexical material.

Still, there is a major difference regarding meaning and function. Prototypical stems and free morphemes have lexical and denotative meaning. (Coulmas 1981, 76) Prototypical routine formulae, on the other hand, have idiomatic and functional meaning, i.e. their meaning is derived from the actual usage and cannot be derived directly from the morphosyntactic and semantic structure of the phrase. In addition, the rules which govern the usage of routine formulae are not only grammatical but also social rules. An analysis and description which is only syntactic-semantic is therefore incomplete. (Coulmas 1981, 16)

On the basis of these arguments, usage of language is defined in terms of semantic frame, semantic script and lexical material in this thesis, in other words, it includes situational context and internal structure. This means that the lexical material here is embedded in a wider context than the issues addressed in the vocabulary section. Thus, it is justified to discuss the lexical material in language usage separately from the rest of the lexicon.

5.1.3.2. Indigenous Genres

As Ega is a spoken language, there is no written literary tradition. Nevertheless, like in other African cultures there is an oratory tradition.

The history of the African oratory tradition goes back to the Egyptian antiquity in terms of oratorical format, styles and dynamics. Its traditions conform to particular African cultural expectations and are considered standard and authoritative. In this line of reasoning African orature has to be analysed within this of approved and recognised set of traditional standards. (Knowles-Borishade 1990, 448)

Orature entertains but it also informs and instructs. (Agatucci n.d.) It sets models for human and cultural behaviour (Scheub 1985, 2) and establishes and reinforces social and political order within African communities (Mutere n.d.), i.e. the oratory tradition is used as an instrument for solving conflicts by negotiation and by setting moral guidelines. In other words, the oratory tradition is a medium for organising, examining and interpreting experiences of an audience. (Scheub 1985, 2)

Scheub (1985, 2) divides African orature into three major genres:

- the riddle and the lyric poem
- the proverb
- the tale, heroic poetry and the epic

The importance of the tale within the framework of African orature is emphasized by Obiechina (1993):

> The story itself is a primary form of the oral tradition, primary as a mode of conveying culture, experience and values and as a means of transmitting knowledge, wisdom, feelings and attitudes in oral societies.

To cut a long story short, indigenous oratory genres are part of language usage and are the cultural heritage of the people which is expressed in language. Therefore indigenous genres are an important constituent of a primer for endangered languages.

5.1.3.3. Classroom language

One context for language development is using the endangered language in actual language classes. This takes the language to a new field where new vocabulary and possibly also new discourse patterns are needed.

Collections of classroom phrases are included in many foreign language textbooks. Even so, I have not come across any systematic approach to setting up lists of phrases needed in class.

One way of compiling up such a list is by pursuing an observation-based bottom-up approach and noting down classroom phrases whenever they occur. The advantage of this approach is that it can start without much theoretical preparation and that it will reflect the actual needs of the learners and teachers. The disadvantage from a linguistic point of view is that this approach lacks systemacy.

Systemacy for the contexts of usage can be achieved by employing the matrix for context description outlined by Hopkins (2002, pp. 206). This matrix requires a description of domains, situations and conditions and constraints on language usage. For the latter, speech acts and interlocutors are relevant categories for classroom usage.

The domain is already specified in the term classroom language – it is the institutional level and more specifically speaking the domain school.

With respect to situations, Neuf-Münkel/Roland (1994, 108) distinguish between simulated communicative exercises which merely practice a new structure and real communicative events in class which take place on a meta level and relate to processes going on in class. Another dimension by which the situation is characterised is the mode the learners are required to use, that is whether it is the perceptive or productive mode. Typical teacher phrases such as *Open your books!* need to be understood by the learners. Other phrases, however, have to be produced by the learners, too, when they want to use their language as a meta language in class. A typical phrase for this is *What does XYZ mean in Ega?*. Furthermore, a situation is characterised by the number of interlocutors, that is whether one, two or more people are actively involved in the discourse. This changes the situation because in situations like discussions learners also need pragmatic competence such as skills relating to turn taking or how to voice an opinion without threatening face.

Speech acts and interlocutors are closely related. With respect to interlocutors, there are the teacher, the learner as an individual and the learners as a group. For each interlocutor there are typical phrases and speech acts. Classroom language includes one-directional imperatives, two-directional questions and answers and multi-directional discussions involving a number of speech acts.

Developing the endangered language so that it can be used as a classroom language is in the interest of all the target groups described above. Language reacquirers may want and, providing that the teachers use the language, need the endangered language in class for actual language learning. For language maintainers and developers who still go to school, the classroom is a new context in which they can use their language and considering the results of the Ega sociolinguistic survey (cf. above) it is likely that among the Ega people there is genuine interest in using Ega in school and in class.

5.2. *Vocabulary or Semantics*

This section is concerned with the question of which lexical items are to be taught. The idea is to set up a lexicon for teaching endangered languages.

5.2.1. Central concepts

A lexicon is defined by Gibbon/Mertins/Moore (2000, 433) as

> a collection of lexical items such as words or idioms organised within a lexical

macrostructure (a list, a tree structure, etc.), each of which has a regular microstructure to which types of lexical information are assigned

Gibbon/Mertins/Moore (2000, 401) treat *lexicon* and *dictionary* as synonyms and define *dictionary* simply as a book which contains a list of lexical entries (usually words) and their properties. In accordance with that, the two terms are used interchangeably here.

In addition to micro- and macrostructure, Gibbon (2003[b]) defines mesostructure as a characteristic feature of dictionaries. The microstructure of a lexicon describes the structure of the individual lexical entries. It includes types of lexical information and the way in which types of lexical information are organised. Mesostructure refers to the relations between lexical entries and other sources of information. These relations can be within the lexicon, e.g. references to a general descriptive section or to other lexical entries, or the references can relate to external sources such as corpora or contexts of usage. Finally, the macrostructure of a lexicon relates to the overall organisation of the dictionary. It includes the lexical entries, their microstructure and the mesostructure. (Gibbon 2003[b])

Various terms are used for the words which enter a lexicon. Gibbon/Mertins/Moore (2000, 432) use *lexical item* in their definition of a lexicon and define this term simply as a unit of language such as a word or an idiom which is inventarised in a lexicon. According to them, *lexical item* and *lexical entry* are synonyms. In addition, they use the same definition for the term *lexical unit*. If there are variant forms of a lexical entry, Gibbon/Mertins/Moore (2003, 431) use *lemma* as a term for the headword or lexical access key under which these variants are listed.

Dictionaries can be organised semasiologically or onomasiologically. Semasiological dictionaries use words as the lexical entry and give information on their meaning. The entries are mostly in alphabetic order. Reichmann (1990, 1063) calls these lexica dictionaries for reception because they help the user to find out about the meaning of a word. Onomasiological dictionaries, on the contrary, use a concept as the entry and give words which are related to this concept. The order in which the concepts are listed varies. Reichmann sees these dictionaries as production dictionaries because they help the reader find word forms. Onomasiological dictionaries are important for teaching vocabulary as they give the user access to a range of lexical items which are all related to one concept or situation. This is helpful when a module on a particular topic is set up.

A headword in an onomasiological dictionary can be characterised as covering a lexical field (cf. *onomasiological capacity* as discussed below). The terms *lexical field* and *semantic field* are often used like synonyms (e.g. Lehrer 1974). Lehrer (1974, 1) defines *lexical field* in a broad sense as "a group of words closely related in meaning, often subsumed under a general term". According to her,

> the analysis of a lexical field includes the relationship of words that contrast paradigmatically (all belong to the same part of speech), and those of other parts of speech that are related morphologically and semantically, for example the relationship between *fly* and *wing* or *tutor*, *tutorial*, and *tuition*. (Lehrer 1974, 22, italics hers)

5.2.2. Issues in vocabulary

This section discusses which words should be included in a lexicon for a primer for endangered languages, i.e. the issue of core vocabulary, and how this lexicon should be organised to suit the needs of teaching endangered languages at primer level.

There are two issues which might be expected in this section but will not be discussed. One issue is word formation. This may be considered relevant in a vocabulary section because this is one way of how words enter a language and it is relevant for deciding whether regularly derived or compounded words should be lemmas in a lexicon. This question will not be answered here because for the language discussed in this paper there is not enough data on word formation to serve as a basis for a decision. Therefore, formal aspects are discussed in the morphology section as derivation is part of morphology and lexicon issues are not discussed at all with respect to word formation.

Also, the question whether routine formulae in the sense of fixed phrases should enter the lexicon is not discussed here but in section 5.1, which focuses on usage.

5.2.2.1. Core Vocabulary

Ideally, language teaching enables the learner to access the whole lexicon of a language both actively and passively. Yet, a complete lexicon is impossible to fix in writing. Even for thoroughly documented languages, a written lexicon does not contain all the possible words. There are always words which can be regularly derived and are therefore not included in the lexicon. Listing all the lexical items in a language is even more difficult if a language has not yet been thoroughly documented and described. In these cases, a comprehensive overview of the inventory of the language is not available. Moreover, in both cases the vocabulary of a language is subject to constant changes.

For a primer, however, this does not necessarily pose too much of a problem. A primer gives an introduction to a language and serves as a starting point. Therefore what is needed for a primer is a set of core vocabulary rather than a full lexicon with all the possible lexical items.

Kühn (1979, 60) defines core vocabulary as a minimal set of vocabulary which provides the user with the lexical equipment needed for linguistic interaction. Setting up a set of core vocabulary requires precise definitions and criteria for its scope and content.

Former criteria for setting up a set of core vocabulary

According to Kühn (1979, pp. 23), former lexicographers have relied on the following criteria for establishing a set of core vocabulary:

- STABILITY: more stable than other vocabulary
- BASIC CONCEPT: words which cover a wide range of meanings and which by themselves are vague in meaning
- BASIC MORPHEME: suitable for compounding and derivation
- NEUTRAL with respect to style
- EXPANDABLE semantically: can extend its range of usage via meaning shift
- is used in idioms and fixed phrases
- indispensable and important for usage
- FREQUENT, although amendments are made with respect to grammatical vs. content words, spoken vs. written language, etc.

These criteria correspond to criteria found in my own survey of students' sets of core vocabulary. I examined five different dictionaries (Litvinov 2000, Blass/Friedrich 1969, Nickolaus 1977, Langenscheidt 1987, Häublein/Jenkins 1998), their prefaces and the criteria they give for the words which are included. The criteria the authors give are similar to the criteria found by Kühn:

- frequency, versatility, part of current colloquial language (Häublein/Jenkins 1998)
- frequency in spoken and written language (Nickolaus 1977)
- frequency and importance (Blass/Friedrich 1969)
- frequency, degree of familiarity, usefulness (Langenscheidt 1987)

Kühn (1979, 40) criticizes that these criteria are hard to operationalise and therefore often subjective rather than objective. Furthermore, they lead to biased conclusions if they are applied to unbalanced, unrepresentative, written language only corpora. In addition, as languages constantly change, they are only valid for a certain period of time.

Improved criteria for setting up a set of core vocabulary

As a consequence of the shortcomings of past criteria, Kühn's own definition of core vocabulary (Kühn 1979, 59) puts less emphasis on perceived statistical objectivity, which he considers impossible to achieve in any case. Instead, he underlines that setting up a core vocabulary is not just a matter of statistics and frequency counts, however elaborated those may be, but instead must be based on people's way of life and their needs in terms of linguistic interaction. This view is shared by the Council of Europe (2002) and it supports the objectives and approach to primer compiling outlined in section 4.1.

Kühn's (1979, pp. 58) criteria for setting up a set of core vocabulary for both first language (L1) and foreign language (L2) teaching are based on language usage and include the components

- interlocutors
- communicative context
- communicative intention and discourse patters
- mode of communication
- topics in communication

Incidentally, Lehmann (2001, 13) proposes almost identical parameters, namely speech act participant, context of speech act, task, topic and channel, as parameters for the selection of speech situations and text genres in language documentation. This emphasizes the close link between language documentation and collecting and preparing primer material for endangered languages.

In Kühn's model, the interlocutors in the communication are the potential users and their needs, which for example are different for children and for adults. For this thesis, primer designers and teachers are added as potential users, as they also need core vocabulary for reference for setting up teaching material.

The context of communication is defined by Kühn as a list of relevant situations the user might get into. These situations make the choice of vocabulary predictable and influence the vocabulary that will be needed. Semantically speaking, I suggest considering these contexts semantic frames.

Communicative intentions and the discourse patterns of the communicative act are summarised in a list of the various illocutionary acts. This list makes it easier to determine the vocabulary which will be needed. Semantically speaking, I suggest considering the discourse patterns semantic scripts. As outlined above, there are the words and phrases related to the semantic script. This leads to the described overlap between vocabulary and usage as discourse patterns involve routine formulae and pragmatic idioms.

For modes of communication, Kühn suggests oral vs. written, direct vs. indirect, bilateral vs. monolateral, instrumental vs. expressive and receptive vs. interactive.

The topics included in a definition of core vocabulary must be topics with communicative relevance. Kühn suggests ordering a meal in a restaurant, asking for directions or describing a problem to a doctor. Of course, these “touristy” topics are less relevant for endangered languages. Therefore, what is required is, as Kühn demands, a careful analysis of the topics in the actual context of usage.

Core vocabulary for endangered languages

Kühn (1979) does not address the issue of teaching endangered languages. Therefore, before his model is applied to this context, it has to be examined whether it is suitable or whether it may have to be adapted.

Teaching endangered languages is not necessarily like the classical L1 or L2 teaching Kühn focuses on. Whereas language development and language revitalisation can be compared to L1 and L2 teaching respectively, things look different for language maintenance and reacquisition. There, learners are in an “in-between” stage because they know too little to be considered first language learners who by and large have the language at their command and just have to systematise their knowledge. Still, they know too much to count as foreign language learners. Also, they are not classical Western European L2 learners because they will not go abroad to speak the language and they will not be in a situation where they are forced to use the L2 like for example tourists in Britain who have to use English or be stuck.

The needs and objectives in teaching endangered languages are mainly the everyday communicative contexts similar to the ones Kühn describes but they are not only that. In addition, the contexts cover

- oratory genres like story telling
- abstract discussions on beliefs or the history of the community
- descriptions and instructions like how to prepare traditional meals such as for Ega chicken in banana leaves

Taken together, this mounts up to more than everyday communicative competence and includes transmission of cultural heritage, history and identity, too.

As a consequence, Kühn's proposals are extended to include two approaches to establishing a set of core vocabulary – a bottom-up, data-driven approach and a top-down theory-driven approach.

The data-driven approach requires a representative corpus and a valid methodology which extracts the relevant words and phrases. The theory-based approach requires a model which states which factors have to be taken into account and which fields and words have to be covered. Also, the model must include reasons for its selection of factors.

In language documentation, a merely data-driven approach is unlikely to work. Documentation is not fully comprehensive description, it cannot guarantee for a representative corpus. For example the Ega language data corpus is not yet representative enough as it lacks for instance every-day communication on the field, one of the major areas of work.

A fully theory-driven approach, however, runs the risk of imposing a ready-made linguistic template which may not necessarily fit. Linguistic models developed so far can maximally suit the languages documented and described so far. It is still unknown what so far undocumented languages can contribute to these models – maybe the indigenous view of the world and its systemacy differ from the existing models and a ready-made model may fail to notice this.

Therefore, as mentioned above, a combination of a systematic theory-based approach and a data-based approach is needed. One possible approach is:

- careful sampling of usage scenarios according to the criteria given by Kühn (1979)
- double-checking this collection against the reality in the community via observation, consultation with speakers etc. and adding the speakers' needs and wishes, thus adapting the model to the community
- filling this matrix of scenarios with material from a corpus and filling the gaps through observation and elicitation
- double-checking by native speakers to ensure that the data is both correct and approved of by the community (e.g. no mistakes, no violation of taboos, no revelation of secrets)

5.2.2.2. Lexical Fields in onomasiological dictionaries

Semasiological dictionaries for Ega already exist (cf. Gbery n.d., Bole-Richard n.d.[c]). Therefore this discussion concentrates on onomasiological dictionaries and the notion of lexical fields.

Like other dictionaries, onomasiological dictionaries can be characterised in terms of their specific macrostructure, microstructure and mesostructure. Reichmann (1990) discusses the first two structures, the third is not mentioned as such but is included in his demands on onomasiological dictionaries. Reichmann's findings are presented here and discussed in terms of relevance for primer designing.

Microstructure

Reichmann (1990, pp. 1057) defines two obligatory fields in onomasiological dictionaries: A basic sign which represents a concept and one or more terms which relate to this basic concept. Based on this, microstructure is defined as an ordered set, or formally speaking an n-tuple, of information given for a basic sign. This definition is similar to Gibbon's (2003[b]).

The tuple contains information such as contexts of usage of terms or expressions in lexical relation to the basic sign. For contexts of usage, Reichmann found collections of phrases, clauses, sentences, fragments, sayings and others. He also encountered semantic frames, which included information for example on discourse patterns, participants, circumstances, conditions, results and objects and concepts related to the situation.

Regarding lexical relations, Reichmann found synonyms and partial synonyms, antonyms and complenyms, hyponyms and hyperonyms, analogies and combinations of all these relations.

In addition to context and lexical relations, Reichmann found entries which also occur in semasiological dictionaries. These included pronunciation, stress placement, spelling, comments on the morphology of inflexion and word formation, etymology, semantic comments, dimensions of usage, encyclopaedic explanations, examples of phrases, clauses, sentences, references to occurrences of the words in literature, information on frequency and finally references to other literature on the subject.

A microstructure as detailed as this requires exhaustive documentation and description, which is often not available for endangered languages. On the other hand, a lexicon for a primer does not require the full set of lexical relations or information on etymology to name only two features. For a usage-based approach it is essential to have information on the context of usage and basic information on lexical relations such as hyponyms, hyperonyms and taxonomies in order to fill lexical fields like for example the lexical fields *greetings* or *colour words*.

Mesostructure

Reichmann (1990, 1064) found alphabetically organised registers as appendices in onomasiological dictionaries which help users find words. In addition, Reichmann demands that onomasiological dictionaries must also contain references within the dictionary itself. (Reichmann 1990, 1063) This goes in line with the idea of a mesostructure as defined above, i.e. mesostructure which links bits of information between lexical entries within the lexicon and lexical entries and external sources of information.

Macrostructure

Reichmann (1990, 1058) understands *macrostructure* as the set of basic concepts. This differs from Gibbon's (2003[a]) definition, who additionally includes microstructure and mesostructure in his concept of macrostructure.

Reichmann (1990, pp. 1058) discusses a number of issues relating to macrostructure and basic signs. Three of them, motivation, selection and order of basic signs, will be discussed here and checked how they can be applied to a primer.

Motivations for basic signs relate to why certain concepts are chosen as basic signs and others are not. Three of the criteria Reichmann mentions are usable for primer making:

- linguistic principles, i.e. semantic relations like synonyms and antonyms
- onomasiological capacity of a concept, i.e. in how far a basic concept is capable of covering a whole lexical or onomasiological field
- pragmatics of usage, i.e. where the anticipated user is most likely to start when s/he wants to look up a word

A non-native linguist has learned how to cope with the first two aspects of dictionary making, although Reichmann qualifies this by stressing that the onomasiological capacity of basic sign must be empirically explored, too. A qualified language teacher is trained to anticipate the needs of primer designers and teachers. Neither a linguist nor a teacher, however, are capable of predicting native speakers' search strategies as they do not necessarily know how the native speakers organise and conceptualise their world. Therefore a combination of a general model-based approach and a specific approach centred on the target language and its speakers is needed again.

The selection of basic signs refers to which types of concepts can serve as basic signs or headwords. Reichmann gives a number of possible types, one rather concrete and one very vague principle can be applied to the selection of basic signs for a primer:

- units in an ideological system of perceiving and explaining reality
- a certain set of terms considered relevant for certain purposes

Units in an ideological system can be used in cases where the native speakers' perception of the reality and their classification of items in the world has been elicited and described.

The second principle becomes useful once the variables have been set and the notion of *relevant* has been operationalised. For example, the phrase *certain purposes* can be defined as situations in everyday communication, the phrase *certain set of terms* can be defined as all the vocabulary which occurs in these situations and *relevant* can be operationalised as a distinction between vocabulary needed by language developers, maintainers, reacquirers or revitalisers.

Independent of which principle is applied, Reichmann (1990, 1063) emphasizes that basic signs and concepts in onomasiological dictionaries must have an empirical basis. Linguistic research has to find these semantic categories or nodes in the language and the culture.

The last issue is the order within the set of basic signs. Reichmann found various systems: ideologically based classifications, alphabetic orders and combinations of both systems. At the same time, he reminds the reader that in pluralist democratic systems there is more than one way of perceiving and explaining the world. There is no generally accepted ideology and as a result it is inevitable to have competing macrostructures and sets of basic signs.

A non-native linguist is unlikely to be familiar with the "ideology" of the target language and culture. Nonetheless, an unrelated ideology must not be imposed on the dictionary either because it is unlikely to match the native system. Alphabetic orders go against the whole idea of an onomasiological, meaning-based dictionary. Moreover, languages which are so far only spoken do not have an officially approved of alphabetic order for their graphemes. As a consequence, the question of how to establish an order has to be left unanswered for the moment.

A survey of existing onomasiologically organised sets of core vocabulary

In order to find organisational principles which can be applied to the organisation of Ega core vocabulary, I examined five onomasiologically organised sets of core vocabulary and their onomasiological macrostructure, i.e. their semantic organisational principles. (Litvinov 2000, Blass/Friedrich 1969, Nickolaus 1977, Langenscheidt 1987, Häublein/Jenkins 1998) The findings do not help very much.

None of the authors explain their underlying model of semantic organisation, which on the other hand is understandable considering that the prefaces are directed at the users of the lexicon and not at a linguist interested in a theoretical model.

All of the dictionaries differ greatly both in terms of hierarchy and number of levels and in terms of number of top nodes. Some distinguish between two levels (headers and sub headers), others have six levels of organisation. In one dictionary, the top level consists of three major clusters (topic-related terms, general terms, miscellaneous), another author uses up to thirty groups.

None of the classifications manage to do without a miscellaneous or generals category.

Most of the authors mix semantic, syntactic/POS and pragmatic categories.

Improving the existing approaches

Regarding the motivation and selection of basic signs, I argue for a combination of pragmatics of usage and semantic relations. For the usage part, the communicative situations suggested by Kühn (1979) and the contexts of usage elicited in the sociolinguistic survey among the Ega are a useful starting point. The communicative situations are well-motivated basic signs because they are theoretically well founded and empirically verified. They consist of situations which are relevant for the people and can be applied easily to teaching material which is based on the same situations. In addition, I suggest using theoretically established lexical fields such as colour words, numbers, body parts etc. as elicited for Ega in the WALDS[6] questionnaire. The theoretically founded approach and the empirically elicited data represent two parallel layers which reflect the two parallel processes in the research and establish both model based and empirically confirmed headwords.

Regarding the order of the basic signs, I suggest a flat structure. I do not think it is

---

[6] The West African Language Data Sheet (Kropp-Dakubu 1980) consists of a wordlist, a list of grammatical phrases and short sentences and meta data on the language and is used in language

necessary to introduce a linear order. Firstly, core vocabulary does not include all the words of a language and a flat structure is considered sufficient to enable access to all the words. Secondly, having all the basic sign at the same level without any prearranged sequence reproduces the idea of modules which will be outlined below in chapter 7.

Regarding the mesostructure, I suggest links between similar topics, re-occurring words and other overlaps and links to theoretically established fields.

In order to avoid mixing semantic, syntactic and pragmatic categories as basic signs I suggest working with a multi-dimensional grid. The grid can either use three dimensions (semantics, syntax, pragmatics) or two dimensions (semantics, syntax). Pragmatic idioms and routine formulae can either be added as the syntactic category *fixed phrases* or as the semantic category *speech acts*. Once the multidimensional grid has been approved of from a native-speakers point of view, it can be filled with words and phrases from the recorded communicative situations, the lexica, the questionnaires and by the language community itself.

Conclusion

There is no widely accepted template for a meaning-based organisation of core vocabulary to use for a primer. Instead there are principles relating to macro-, meso- and microstructure. All of them heavily depend on the objectives for the vocabulary collection, the intended usage and the users. A possible way to go about is:

- State the objectives, intended usage and intended users.
- On macrostructure:
    - Decide on lexical fields to be included.
    - Decide on their structure (e.g. ideological vs. alphabetical vs. combined, tree vs. flat structure).
    - Order the lexical fields.
- On microstructure:
    - Decide on the categories to be included (e.g. usage, semantic, other linguistic information).
    - Add the information.
- On mesostructure:
    - Find overlaps between the lexical entries and refer to them.

---

documentation.

- Add references to other parts of the lexicon (e.g. a grammar, a pronunciation guide).
- Add references to external data.

5.3. Syntax

5.3.1. Central concepts

Syntax is understood here as a system which allows speakers to construct and likewise hearers to parse formally correct phrases and sentences. The focus on the phrase and sentence level is because the word level has already been dealt with in the morphology section and the text level will be dealt with in the usage section.

Linguistic documentation and description has to give answers to a number of central questions in order to present a basis for a primer. For phrases, these questions are:

- Which types of phrases occur in the language?
- Which obligatory and which facultative constituents are there in the phrases?
- Which dependencies are there between constituents?
- In which order do the constituents turn up?

On the sentence level, the questions are:

- Which sentences types are there in the language (e.g. declarative, interrogative, imperative)?
- Which dependencies are there within the sentence and between the phrases (e.g. verbal government, agreement)?
- In which default and modified orders do the constituents occur?

If possible, the documentation should provide phrase and sentence templates, which can then be used for primer design and in the actual teaching. (cf. the Ega greeting template in section 6.1.1)

Once these questions have been answered, the syntactic inventory can be divided into two subsets. One subset consists of the syntactic structures which are relevant for a primer, i.e. occur in the functional usage of the language which forms the basis for the primer. The other subset consists of the structures which do not occur in the primer or are, on a linguistic basis, considered more complex and may therefore be introduced at a later stage for advanced learners.

5.3.2. Objectives in teaching

The general objective in syntax is to enable beginners, i.e. language reacquirers, to understand and produce the structures included in the basic subset of the syntactic inventory. Language maintainers and language developers, on the other hand, may be interested in the advanced learners' subset.

5.4. Morphology

5.4.1. Central concepts

Morphology is the subdiscipline of linguistics which studies grammatical units (morphemes and lexemes) and grammatical categories. (Bubenik 1999, vi)

Morphology deals with morphemes, which are the smallest meaningful elements in a language, parts of speech (primary grammatical categories) and their respective secondary grammatical categories. Secondary grammatical categories are for example number for nouns, tense for verbs or case for adjectives. In addition, morphology investigates morphological processes such as inflection and derivation.

Inflection is a process which involves a change in the form of a word to express its relation to other words in the sentence. Inflection does not change the primary grammatical category. (Bubenik 1999, pp. 52) In order to describe and explain inflection, three more terms need to be defined: root, stem and affix. A root is a non-affix lexical content morpheme which cannot be analysed into smaller parts. Roots can be bound or free. A stem is a root morpheme combined with one or more affix morphemes. By adding an additional affix, a new stem and a new word is formed. An affix is a bound morpheme attached to a stem or root morpheme. (Fromkin/Rodman 1998, 75 and 519)

Inflectional morphology is concerned with the secondary grammatical features mentioned above. For nouns, Bubenik (1999, 71), lists the categories number, gender/nominal class, case, agreement/alignment. Tallermann (1998, 52) adds definiteness to this list. For verbs, Bubenik (1999, 71) lists person, number, tense, aspect, mood, voice. Tallermann (1998, 57) adds agreement to this list[7].

Derivation involves processes whereby new words are formed from existing words or bases. Derivation usually changes the primary grammatical category, that is it derives one part of speech from another. (Bubenik 1999, pp. 52)

A derivate (or derived/complex lexeme) is a lexeme whose stem is formed from a simpler stem, which is the derivational base, by some kind of morphological modification, most commonly affixation. (Bubenik 1999, 168)

Derivation and compounding are the two processes involved in word formation. Word formation is the branch of linguistics which studies the patterns by which a language forms new lexical units, i.e. new words. A compound is defined as a lexeme whose stem is formed by combining two or more stems which may be separated by an interfix. (Bubenik 1999, pp. 167)

Regarding a primer for endangered languages, the documentary and descriptive morphological basis must provide answer to the following questions:

- Which primary grammatical categories does the language use?
- Which secondary grammatical categories are associated with the primary grammatical categories?
- How are secondary grammatical categories expressed in the language? (e.g. roots, stems, affixes, tone)
- Which inflectional processes occur in the language for which primary grammatical categories?
- Which word formation processes occur in the language?

For objectives in teaching morphology, compare objectives in teaching syntax.

5.5. Phonetics and phonology

5.5.1. Central concepts

Phonetics studies the articulatory, acoustic and auditory characteristics of speech, whereas phonology is concerned with the relations among speech sounds in particular languages. (Trask 1996, pp. 270). In this view, phonetics is more concerned with the formal aspects of speech and contrasts with phonology which focuses on functional aspects of speech.

Phonology distinguishes between two levels – a segmental and a suprasegmental or prosodic domain. Phonemes and tonemes, i.e. consonants, vowels, lexical and grammatical tone, are located in the segmental domain. The suprasegmental domain

---

[7] For detailed guidelines on what may occur in a language consult Payne (1997).

includes for example stress and intonation. (Trask 1996, 343)

Tone is the phonological equivalent to the phonetic concept of pitch. As that, it may also be associated with the suprasegmental domain. However, the term *suprasegmental* is defined by Trask (1996, 343) as

> pertaining to a phonological element whose domain is something larger than a single segment and whose phonetic realisation can only be described by reference to adjoining domains in the same utterance

A lexical and grammatical tone is aligned with a specific vowel, i.e. a single segment, and is not dependent on a context in its default citation form. It may be affected by processes such as declination or downstep when the word is embedded in a context but this is a different question. Therefore it seems legitimate to see lexical and grammatical tone as part of the segmental level.

In addition to segmental and prosodic issues, a primer must also deal with the processes speech production and speech perception. Both processes play are role in language acquisition and it is not yet clear which process facilitates the other. There is evidence and counter-evidence for both directions and especially for the acquisition of lexical tone the underlying processes are far from clear. Therefore both processes must be included in the primer.

5.5.2. Objectives

The specific objectives for phonology differ depending on the target group. Those aiming at language development or maintenance will already be fluent and competent in terms of phonology so that they do not have to be included in the teaching of production.

For perception, things look different. There, speakers have to learn to discriminate between phonemes and between tonemes and to identify the phonemes and tonemes because this is a prerequisite for being able to read and write in a phonemically based orthography.

Learners interested in language reacquisition will be in a typical second language acquisition situation. They have to acquire skills in perception and production on both articulation, i.e. on the segmental level, and intonation, i.e. on the suprasegmental level.

5.6. Orthography and orthography development

A major prerequisite for a written primer for an endangered language is that there is a writing system, a script and an orthography. Therefore the following section discusses what has to be taken into consideration when a new orthography is developed for a so far unwritten language.

5.6.1. Central Concepts

Coulmas (2003, 35) distinguishes between writing system, script and orthography. He defines *writing system* as follows

> a set of signs used to represent units of language with the purpose of recording messages [...] as well as the general principles of their employment (Coulmas 1999, 137)

In an abstract conceptualisation of writing system (Coulmas 2003, 35), the term relates to a structural typology of different ways of writing which includes

- logographic or word writing systems
- syllabic writing systems
- phonetic writing systems
- variant forms of these systems

The basic functional units in a writing system in a very general sense are called *signs* by Coulmas (2003, 35). For Semitic-derived writing systems, Coulmas also uses the term *letter*. One or more letters may combine into a grapheme. A grapheme is an abstract type of letter and which has a specific function in a given writing system. Graphemic units can consist of individual letters, i.e. be monographs, or they can be digraphs, trigraphs, and so on. (Coulmas 2003, 99)

Script is a term for the graphic form of the units of a writing system. (Coulmas 2003, 35) Coulmas does not explicitly comment on the relation between the graphemes and the script. I do, however, see a mutual dependency there: The script determines which set of letters to use (i.e. the Roman or the Cyrillic script), it relates to the form. The graphemes add meaning and structure to the letters or combination of letters. In a way this can be explained in the above mentioned textlinguistic model of documents which is adapted here to characterise the interdependences within writing systems:

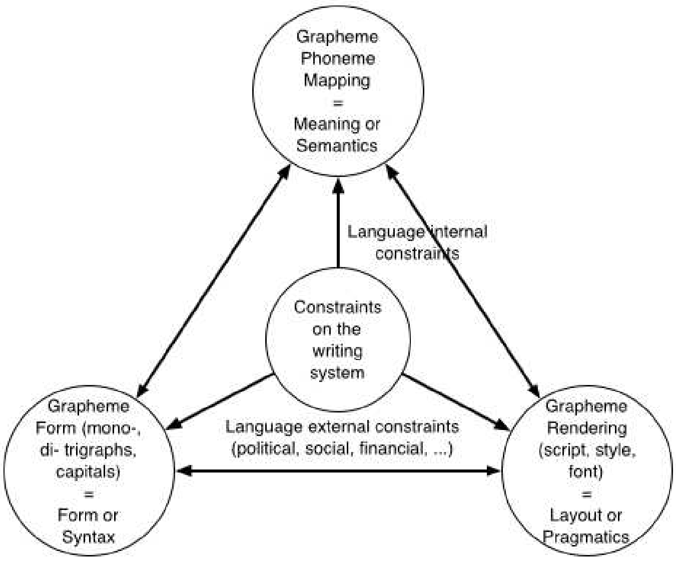


Figure 7: Linguistic interdependences within a writing system

Constraints affecting the choices to be made in orthography development are discussed by Coulmas (1999), although he does not put them under the notion of constraints. They are outlined in more detail below.

The term orthography refers to the normative spelling conventions of a language. An orthography consists of a word list and orthographic rules, which are normative rules of regulation and determine particular aspects such as sound-letter correspondences, letter sequencing, capitalisation, spacing, word segmentation or punctuation. (Coulmas 1999, pp. 137)

Writing systems and orthographies are linked in the way that the "orthography makes a selection of the possibilities offered by a writing system for recording a particular language in a uniform and standardised way". (Coulmas 1999, 137)

5.6.2. Objectives

For all the other levels of linguistic analysis and language teaching the linguistic structures already exist in the language. They just have to be documented, analysed and described and then they can be taught. This is different for teaching to read and write in a language which has so far only been used in the spoken modality. Here, the linguistic structures have to be developed first. Therefore this section distinguishes between objectives in orthography development and objectives in teaching orthography.

For new orthographies for endangered languages, developers need to decide on

- the writing system

- the script, graphemes and letters
- the orthographic rules

while considering a number of constraints relating for example to the linguistic structure of the language or to social and political aspects.

If the language uses digraphs and trigraphs, it is important to explicitly teach and practice these combinations, too, so that readers are able to detect and decode them and will not mix them up with sequences of monographs. (Feldmeier 2003, 28)

A crucial point in literacy development is translating the perception of fluent speech into abstract phonemic and tonemic categories and then into graphemes. This is something learners may not be used to. In addition, learners which are literate in one type of writing system, e.g. a mixed system using logographic, morphemic and phonemic elements such as English (Coulmas 2003, 188), are also only partly used to this translation activity. Ladefoged (2003, 81) reports on an occasion which nicely illustrates this problem. Ladefoged worked with a Nigerian student who was a native speaker of a tone language and was asked whether a certain word had a high or a mid tone. Apparently the student was not able to translate his perception into the tonal category and so he replied: “Don’t ask me. All I know is that one moment you were talking about a kind of worm and the next asking me about dinner.”

A way of solving this translation problem is designing a prep course for a primer which includes both phonology and orthography and explicitly relates the two modalities of speaking and writing to each other. In addition to this prep course it is possible to develop supplementary material such as letter-picture charts for reference. This may raise awareness of the notion of phonemic and tonemic categories on the one hand and graphemes on the other hand and may aid in encoding spoken language into graphic categories and decoding spoken language from graphic representations.

To sum up, the contents for teaching an orthography in a functional, pragmatic approach require less meta information and more skills. Objectives suggested here are to teach skills relating to understanding and applying

- the relations between written and spoken language
- the relation between graphemes and phonemes, graphemes and tonemes, graphemes and words and finally words and concepts
- which graphemes there are
- orthographic rules

5.6.3. Issues in orthography development

5.6.3.1. Aspects in orthography development

Social aspects

Coulmas (1999, pp. 138) identifies three social aspects which have to be taken into account when new orthographies are developed: politics, religion and the question of identity.

Politics may have an effect on the use of a writing system when there are changes in the political direction of the society. This happened for example in Turkey with the adoption of the Roman alphabet or after the disintegration of the USSR, when the Cyrillic alphabet was dropped in favour of the Arabic alphabet in some of the former republics.

Religion influences the choice of writing system in the way that Islamic societies tend to use the Arabic alphabet, whereas Christian societies usually choose the Roman alphabet.

Sometimes an orthography is used as an identity marker in order to either set oneself apart from another society that the society does not want to be associated with (German, Dutch, Belgian Dutch, Swiss German, etc.) or in order to be closer to a prestige language, which often is a former colonial language or an official language.

Structural aspects

With respect to writing systems, there is the choice between phonemic, syllabic or orthographic systems or mixed systems.

Coulmas (1999, 140) recommends using the basic principle of phonemic characters, where units of speech sound are isolated and used as referents of graphic signs. He claims that this principle has over the times proven to be a sensible approach but does not present further support for this claim.

For languages with relatively straightforward phonotactics, e.g. with only V and CV syllable structures, syllabic alphabets may be an alternative worth considering. If these languages have only a restricted and relatively small number of possible syllables, it is possible and manageable to have one sign for each syllable.

Finally, logographic characters are also likely to enter the writing system. Many languages use Arabic numbers representing numerals, i.e. words, and there is punctuation which for example indicates speech acts, i.e. various functional types at the sentence level.

In addition, morphological information is encoded in orthographic rules, for example when proper names are capitalised.

Concerning scripts, there is the choice between adopting an existing approach and, if necessary augmenting it when it does not suit the needs of the language or designing a new script. For existing scripts, an alphabet worth mentioning is the *Africa alphabet*. This is a writing system which consists of eight vowels, twenty-seven consonants and one diacritic element. It was designed by the International Institute of African Languages and Cultures in London in 1930. The intention was to augment and extend the Latin alphabet in order to adapt it to the phonemic inventories of African languages. (Coulmas 2003, 102) This is aiming high, especially when considering that there are African languages which have a consonant inventory of over 90 sounds such as in the Namibian language !XŨ. (Crystal 1987, 168) Thus, this script is doomed to fail.

5.6.3.2. Constraints on orthography development

Among his discussion of structural aspects in orthography development, Coulmas (1999, 140) mentions two constraints, namely economy and precision on the one hand and transferability on the other. Regarding the former, he states that the orthography should not go into every detail but should "strike a balance between the requirements of economy and precision" so that it makes reading and writing easy for the user. He does, however, not state how to operationalise "easy" or "hard" to read and write.

Regarding transferability, Coulmas sees difficulties because especially for minorities their language is not the only language spoken by the inhabitants of the area or country. Therefore, Coulmas recommends developing an orthography which ties in with other regional languages or an official language which is used in school:

> By using the writing system of a dominant language as a model for the orthography to be designed for a minority language spoken in its area, the minority potentially benefits from transferability. (Coulmas 1999, 140)

Similar and additional principles have been set up by the Institut de Linguistique Appliquée of Abidjan University (ILA 1996, pp. 4). Together with SIL (Société Internationale de Linguistique), ILA aims at developing writing systems for so far unwritten languages which are both scientific and practical in their nature and fit all the languages spoken in the Ivory Coast. Their principles are:

1. A harmonisation of the writing systems of the four major language families of the Ivorian languages.

2. A writing system that is in conformity with the phonological system of the given language and follows the principle "one sound – one symbol" and where the letter do not map onto different phonemes in the various languages.

3. A writing system that uses standard characters which can be found on any keyboard.

4. A choice of symbols which are close to French to enable learners to progress to French

In the practical application of the constraints, it is very likely that constraints will come in conflict with each other. For example, if only ASCII characters are used but the vowel inventory of the language consists of more than five vowels, then this either means having to violate the ILA principle of mapping one phoneme onto one grapheme or having to violate the ASCII only principle which as a consequence may violate the idea of transferability. Therefore orders of constraints have to be set up, too, which state the relative importance of s constraint compared to other constraints.

5.6.3.3. Steps in developing orthographies

Coulmas (1999, pp. 140) gives a short overview of how a phonemic orthography is designed.

As a prerequisite, Coulmas demands a phonological analysis of the sound system of the language. This essentially precedes the design of a new orthography because without knowing which phonemes there are it is impossible to decide on the number and kind of graphemes to use.

Secondly, Coulmas recommends taking an established writing system (such as the Roman or Arabic alphabet) and then mapping one phoneme onto one grapheme.

Thirdly, if the writing system does not fit the phonological system of language in question, the original system is augmented and modified. These adaptations can be designing digraphs, trigraphs or by employing other complex signs like diacritics.

As Coulmas's comments are only intended to be a short overview, it is hard to criticize them for not going into enough detail. Still, even for a short outline it lacks one crucial element, which is evaluation. Coulmas does discuss orthography evaluation, but apparently does not see it as an integral component of orthography development.

Contrary to this, an evaluation phase which comes into effect after every step is suggested here. This evaluation takes the aspects and constraints discussed above and applies them to the orthography proposal. If the constraints are met, it is justified to go on to the next step. If the constraints do not match the proposal, the step taken before needs to be repeated to adapt the proposal.

As an operationalisation of the concept of a “good” writing system, Coulmas (1999, 138) suggests regarding social aspects, i.e. that the system is accepted by the society and thus has social validity. For structural aspects, a writing system is “good” if it does the structures of the language justice. Unfortunately, the latter point remains rather vague.

5.6.3.4. Criteria for orthography development for endangered languages

Taken together, the various proposals and their criticism result in the following criteria for orthography development:

Social acceptance: The new system takes social aspects such as politics, religion and the question of identity into account. (cf. Coulmas 1999, pp. 138)

Precision and Economy: The writing system and the orthography include as many details as necessary for both native speakers and learners of the languages as a foreign language. All the meaning-distinguishing features (for Ega both phonemes and tones) are represented in a way that they can be identified unambiguously. The balance between economy and precision Coulmas (1999, p. 140) calls for is shifted in favour of precision because language reacquirers may not have a native context in which they learn the language and therefore need additional support to make up for this disadvantage. (cf. section 9.1.3.2)

Phonological accuracy: The writing system is in conformity with the phonological system of the language. It follows the principle “one sound – one symbol” so that there are not ambiguous graphemes. (cf. Coulmas 1999, 140; ILA 1996, pp. 4)

Transferability: Transferability relates to both other regional languages and a possible official language which may be considerably different in terms of structure and genetic affiliation.

The new orthography makes it possible for learners who are already literate in one language to use their skills and knowledge and transfer it to the newly developed writing system. Likewise, it enables learners who start their literary career with the new writing system to progress more smoothly to already established systems for other languages.

This principle combines Coulmas's and ILA's ideas

- that a linguistic minority can benefit from a transfer of knowledge when they learn the official language (Coulmas 1999, 140)
- on the principle of a harmonisation within the spelling conventions for all the Ivorian languages (ILA 1996, pp. 4)

ASCII Characters: The new orthography uses standard keyboard characters (ASCII) wherever possible. This makes publications in the language cheaper and, as a consequence, more likely. Additionally, using the well-established ASCII characters contributes to transferability. (cf. ILA 1996, pp. 4).

Learnability and practicality: The newly developed orthography makes reading and writing easy to learn and easy to use for the people. (cf. Coulmas 1999, 140)

5.7. Summary of general principles for primer-making

In the last sections, a theoretical frame of reference has been developed which looked at specific areas in linguistics and considered relevant issues in making primers for endangered languages. This section sums up these issues and integrates them into a general paradigm of approaches to primer making.

As said above, Himmelmann (1998, 177) calls for a two-way approach to documentary linguistics. This call is taken up here and applied to primer making. For this, Himmelmann's two-way approach is extended and adapted.

The general principles for primer making are integration and co-operation on various levels, e.g. between people, research paradigms, teaching traditions, etc. This principle is sub-divided into more specific characteristics, which leads to a combination of approaches such as this one:

Table 10: Approaches to primer making

Approach A	Approach B
analytic	holistic and synthetic
model-based	observation-based
top-down	bottom-up
theory-driven	data-driven
theory-testing	theory-generating
more controlled	more explorative
general	language-specific
by linguists	by linguists and community

This dichotomy of research paradigms also exists in related empiric disciplines, e.g. in language teaching research (cf. Edmondson/House 2000, 39).

Both approaches have their disadvantages and their advantages. A systematic model-based approach establishes clear guidelines for data collection, description and application. With this approach, gaps in the data are less likely. Still, it cannot be guaranteed that the ready-made linguistic model is appropriate for describing the language. A so far undocumented language may bring forward unknown structures which are not yet included in any model. Hence, the model cannot detect and classify these structures.

An observation-based model is not limited by a linguistic matrix so that this approach is more likely to notice structures beyond the scope of a model. On the other hand, the collected data may lack systemacy. Moreover, it is more difficult to apply quality criteria like objectivity, representativeness or reliability to the data collection.

Combining the approaches means making better use of the advantages and being less restricted by the limits of the individual approaches.

Parallel to the integration of approaches for primer making, teaching traditions may also be combined. For Ega, traditional ways of passing on the language are not sufficient anymore to guarantee the survival of the language. This is clearly shown in the Ega sociolinguistic survey, which found a substantial shift to neighbouring languages among children. (Connell/Ahoua/Gibbon 2001) One way of reversing this shift could be institutional teaching of the language. However, this should not result in imposing an alien, e.g. French, system onto Ega schools. Instead language teaching should start from the system which exists already and combine institutional teaching with traditional methods. This follows the principles laid out in the objectives and for the pragmatic foundation of the primer.

The combined approach runs parallel to Gibbon/Connell/Ahoua/Urua's (2001, 6) demand to involve the community in the documentation in order to guarantee the success of the initiative. Community involvement may for example include:

- curriculum design and development
- language teachers who are native speakers
- cultural heritage teachers who are native speakers, e.g. storytellers, village chief
- native teachers for traditional skills, e.g. vocational training for traditional crafts

6. CONTENT – PRIMER-RELEVANT EGA

In the last chapter, the theoretical outline of a linguistic model for a primer was given. In practice, however, it is not always possible for a number of reasons to put all the theoretical demands on an ideal primer into practice. The language may be extinct already, not all documentary or descriptive data may be accessible or data specially compiled for the primer cannot be gathered because native speakers cannot be contacted. This is exactly what is the case with Ega at the moment. Because of political instability and lack of funding it is currently not possible to do fieldwork in the Ivory Coast among the Ega. As a consequence, data gathered for other purposes has to be used and adapted to the needs of the primer.

This sections presents exemplars from the Ega language documentation and description which can be used for compiling an Ega primer.

6.1. Pragmatics/Usage in Ega

6.1.1. Greeting in Ega

Ega greetings have not yet been documented and described in detail. There is, however, some audio and video data and some observational evidence from a fieldwork visit in an Ega village on the basis of which a description is attempted for the purpose of this primer. This description follows the framework of functional usage of language and discourse introduced in section 9.1. This framework consists of the components semantic frame, semantic script and lexical material.

With respect to semantic frame, Ega greeting formulae differ in terms of

- time of the day
- number of people being greeted
- register, i.e. formal vs. informal

The semantic script is rather straightforward: One person greets another person or group of people and she/he/they greet back. The formulae used for greeting and greeting back are identical. It has not been elicited which person is meant to greet first. After greeting, it is possible ask “How are you?”, the exact procedure for that has not yet been elicited. Table 11 gives an overview over greeting formulae which have been elicited from and confirmed by native speakers. Word segmentation follows the transcript of the native speaker.

Table 11: Overview over Ega greeting formulae[8]

Time of the day	Sunrise	Daytime	Evening/Night
Informal singular	O´flI- mU´ka`	mU´ka`	o´Jo-d<o- mU´ka`
Informal plural	O´flI- mU´ka`nU´	mU´ka`nU´	o´Jo-d<o- mU´ka`nU´
Formal singular	(not elicited)	mU´ka` bE-	(not elicited)
Formal plural	(not elicited)	mU´ka`nU´ bE-	(not elicited)

It appears that the basic greeting formula is *mU´ka`*. In the morning and in the evening it can be supplemented by *O´flI-* or *o´Jo-d<o-* respectively which precede *mU´ka`*. If two or more people are addressed, the plural marker *nU´* which follows *mU´ka`* is used. The politeness marker *bE-* is used for formal greetings. It follows either *mU´ka`* or the plural marker *nU´* if this occurs. In addition to lexical markers for politeness, the formal register or the polite version is also expressed by gestures such as bowing slightly and holding your hands in a particular position. (Bazé Lucien, 2002, personal communication)

Taken together, this results in the following greeting formula template:

8 For audio files of the Ega greetings see *Survival Ega* by Lucien/Salffner (2002).

Table 12: Ega greeting formulae template[9]

	Time	Root	Number	Register
Example	O´flI-	mU´ka`	nU´	bE-
Gloss	morning	hello	PL	formal

NOTE: The gloss in this template is based on speculation and has been set up on the basis of the formulae which have been elicited. It still needs to be confirmed by a native speaker.

6.1.2. Indigenous Genres

The Bielefeld Ega corpus does not contain any riddles. It contains a collection of proverbs transcribed and translated by a native speaker and which has been recorded as a list of proverbs, however not in its actual context of usage. Recordings (both video and audio) of actual contexts of usage have been made for story telling sessions. Some of these stories have been transcribed, annotated and translated. One story has been described and analysed with respect to features of African orature and ritualised communication. This story is used here for further theoretical and practical considerations relating to primer material for endangered languages. The discussion is based on a multi-level linguistic annotation and the description of the video recording of this Ega story[10].

6.1.2.1. Contents for teaching story telling

Parallel to the Ega greetings, Ega story telling is also described here in terms of semantic frame, semantic script and lexical material. Unless otherwise indicated, the description is based on Rafferty/Salffner (2002, pp. 35)

The semantic frame or the situational context of the story telling defines who has to be present to tell a story, i.e. linguistically speaking how many and which interlocutors are needed, and where and when stories are told or not told.

For the interlocutors, Knowles-Borishade (1991, 490) suggests five types of participants which are essential for African story telling: the caller plus chorus, the responders, the spiritual entities, Nommo (a mystical concept for the spoken word) and spiritual harmony. In a questionnaire investigating the Ega explanations and the Ega meta

[9] Conventions for glossing follow Lehmann's (1982) directions for interlinear morphemic translations.
[10] Both the video file and the annotation are included in the appendix.

knowledge on story telling, the village storyteller had a more pragmatic approach to interlocutors. According to him, it takes two people to tell a story – one who tells the story and one who listens to it.

The time of the day when stories are told is fixed rather strictly. Stories must be told either in the evening or at night after all the work is done[11]. If somebody tells a story at daytime, the punishment is gruesome – it is said that then this person's parents will die. The place where stories are told is not fixed, also there is no special sitting order. Everybody sits wherever they want to, which is usually where friends are sitting already.

Script-wise, the story telling is structured like a dialogue where the listener is required to contribute to the story telling, too. The listener is obliged to give feedback to the teller by saying *sEsE*, which means *I listen to you*, every once in a while. This phrase is very important. When the listener does not say *sEsE*, this is taken as a sign that s/he is not listening and the storyteller reminds the listener to pay attention.

The importance of this feedback is nicely illustrated by an episode which happened during a recording session in an Ega village in March 2002. The person who was supposed to say *sEsE* forgot to do so and after a while the storyteller was so irritated by this that he interrupted his story, turned towards her and let off a cascade of words. All that was understandable for non-natives was the occasional frustrated *sEsE*. When the storyteller had finished and picked up his story again and the woman actually said *sEsE*, the storyteller, quite relieved, said *violà* and carried on with this story.

Regarding the sequence of components in a story, a tale may start with a song or a proverb or it may start just like that. There are introduction phrases equivalent to the English *Once upon a time* which are usually used to start a story. Usually a story ends in a moral. The story described here additionally includes a sequence of spoken parts and singing which alternate.

Lexical material is understood in both a narrow and a broad sense. Whereas the narrow sense includes fixed phrases, the broad sense contains visual, acoustic and motional elements.

Regarding the phrases, learners have to be introduced to the feedback phrase *sEsE* and to the introductory phrases. In Ega there are two of these phrases – *mo waa*, which

---

[11] The story telling session recorded in the video took place in the afternoon. This was an exception because in the village there is no electricity and therefore no light at night so the recording had to be done

means *one says*, and *no waa*, which means *I say*. *Mo waa* is used when the storyteller has only known the story for a little while and has only heard about the incident the story is about. Thus, it is equivalent to *Once upon a time*. However, once the storyteller has known the story for some time or knows about the incident it is about first hand, s/he says *no waa* instead.

6.2. Ega vocabulary/semantics

It has been suggested above to organise Ega vocabulary in terms of communicative situations on the one hand and lexical fields as elicited in the WALDS questionnaire on the other hand. So far there is not enough documentation relating to communicative situations, therefore only the lexical fields from the WALDS data are included in this outlook on Ega vocabulary.

6.2.1. Ega lexical fields

The Ega core vocabulary so far is organised along two dimensions: semantic categories, which constitute lexical fields and are organised in a flat structure, and syntactic categories, which are nouns, verbs, adjectives and a category for other parts of speech.
The microstructure of the entries so far consists of the lemma, which is a phonological representation of the word, and a French gloss. For nouns, the plural form is given, too.
In the long run, more information should be added. For the moment, however, this is not possible as the current lexica and the WALDS transcripts are not detailed enough. Also, references and links must be added at a later stage in cases where an entry fits more than one lexical field.
The entries in the Ega core vocabulary come from the wordlist of the Ega WALDS questionnaire. The lexical entries are organised as follows:

- lexical entry; French gloss
- lexical entry/plural; French gloss.

French has been chosen as the gloss language because French is the official language in the Ivory Coast and is understood by many Ega. The grid containing the vocabulary is included in the appendix.

---

at daytime.

6.2.2. Ega word formation

Ega compounding or derivation has hardly been described. Regarding derivation, Gibbon/Ahoua/Salffner (2003) have sketched deverbal nominalization. According to them, nouns are derived from verbs by adding the prefix *E*. *E* stands for an archiphoneme or morphophoneme and is realised as either /e/ or /E/. This is due to the process of vowel harmony (cf. above) and depends on the root of the respective lexeme.
Bole-Richard (1982, pp. 391) gives a few examples for noun compounds consisting of two nouns and states that

- the qualifying lexeme is in first position, the qualified lexeme follows
- each lexeme keeps its own vowel harmony, i.e. the ATR-values within the compound may differ

He does not explain the patterns behind vowel alternations, vowel deletion and tone distribution over the remaining vowels which occur in his examples.
In addition to two nouns as lexemes, Bole-Richard gives an example for a compound consisting of an adjective and a noun but does not say anything else than that this is a compound.

6.3. Ega syntax

Ega syntax has not yet been described in detail. Because of this and as this thesis has to be selective, the focus in syntax will be on noun phrases and declarative sentences.

6.3.1. Noun phrases

6.3.1.1. Basic noun phrase template

Gibbon/Ahoua/Salffner (2003) propose a basic nominal construction template which includes three kinds of modifiers: adjectives, determiners and quantifiers. In addition, they include a slot for an interrogative. The template looks as follows:

Table 13: Ega noun phrase template

Segment	(Interrog)	Prefix	Stem	(Adj)	(Det)	(Quant)
Example		E-	wla´	fU`fU`	ma´	E`JO`
Gloss		CL*n*&PL	hen	white	DEM&PROX	two
English		these two white hens				

At least in English, a fourth modifier is possible, as for example in the phrase *some of these ten white hens*. It has not been investigated whether a construction like this exists in Ega. Still, it is worth exploring whether it exists and where a potential fourth modifier would be inserted.

6.3.1.2. Modifiers I – determiners

Nouns in Ega can be modified with respect to definiteness. According to Bole-Richard, this is done by a suffix. According to Gibbon/Ahoua/Salffner (2003), these modifiers are free morphemes and are considered determiners.

Table 14: Ega determiners

Suffix/DET	Function	Example	Gloss
no suffix/DET	generic use, indefiniteness	e`do- i`do-	a cow cows
tA´ tE´	definiteness	e`do- te´ i`do- te´	the cow the cows
mA´ mE´	demonstrative proximate to the deictic centre	e`do- me´ i`do- fU`fU` ma´	this cow these white cows
NU´	demonstrative remote from the deictic centre	(no example given)	

Considering the example *i`do- fU`fU` ma´* meaning *these white cows*, the determiner approach is more likely. The adjective is inserted between the root and the morpheme indicating the definite demonstrative. If *ma´* were a suffix, this would imply that the adjective *fU`fU`* is an infix. This is possible but unlikely, so far Ega has not been described as using infixes.

6.3.1.3. Modifiers II – numerals

Like Ega nouns, Ega numerals consist of a root and a prefix. Within a noun phrase, ordinal numbers agree with the noun root with respect to the prefix morphophoneme. The phonetic realisation of the archiphoneme depends on the ATR-value of the quantifier root.

Table 15: Agreement between noun and numeral

Ega	Agreement	English
o-te- o`lo-gbo´	o – o	a tree
U`g<lO` u`lo-gbo´	U – u	a village
e`fi` E`JO`	e – E	two children
a-nu- a`JO`	a – a	two heads
I-sI- I`JO`	I – I	two women

There are two exceptions to this agreement.

The first exception relates to zero prefix nouns. The prefix for a numeral is *o`* when the numeral agrees with a singular zero prefix noun. When the noun is a plural zero prefix noun, there is the prefix *pU`a`* for the numeral, which is not an archiphoneme. The example Bole-Richard (1982, 382) gives for this is *two mice – bu`cre´ pU`a`JO`*.

The second exception relates to the *I*-noun class. If a noun of this class occurs in singular, then the prefix is not *i`* but *le`*. Thus, *one head* is realised as *i-nu- le`le-gbo´*.

6.3.1.4. Modifiers III – adjectives

If an adjective is used within a noun phrase to qualify the noun, it does not get a prefix. Bole-Richard (1982, 380) illustrates this with the following example:

Table 16: Ega modified noun phrases I

Ega	i`do-	fU`fU`	ma´
Gloss	CL*n*&PL	white	DEM&PROX
English	these white cows		

Adjectives in predicative use do get a prefix in a similar way as the numerals. As part of the predicate, adjectives consist of a root morpheme and a prefix morpheme. The prefix is one of the archiphonemes and agrees with the prefix of the noun qualified by the adjective. The phonetic realisation of the archiphoneme depends on the ATR-value of the adjective root. The contrasting example to the one above is:

Table 17: Ega modified noun phrases II

Ega	i`do-	me´	wi-	I´fU`fU`
Gloss	CL*n*&PL-cow	DEM&PROX	are	CLSF-white
English	These cows are white.			

In the attributive use, the Ega adjective is *fU`fU`* whereas in the predicative use it is *I´fU`fU`*. Bole-Richard does not explicitly comment on that, he only gives the examples. What is striking with about example is that in Bole-Richard's examples *me´* is almost always used with singular nouns and *ma´* is used with plural nouns. Only this example does not follow this pattern. Either this is a mistake in the data or *me´* can be used with both numbers.

6.3.2. Declarative sentences

The default basic constituent order in declarative sentences in Ega is SVO. The object may be followed by adverbials. (Bole-Richard 1982, 376)

Table 18: Ega word order in declarative sentences

Ega	gi-ge´	pE-	E-na-	a`da`´bla´
Constituents	SBJ	VERB	OBJ	ADV
English	Guigué has bought meat on the market.			

6.4. Ega morphology

Ega morphology has not yet been described in detail. This outline of Ega nominal morphology is based on Gibbon/Ahoua/Salffner (2003) and Bole-Richard (1982). Examples are taken from Bole-Richard (1982).
In this paper, the focus in morphology is on nouns. The relevant secondary categories for nouns are number and class.

6.4.1. Morphology of nouns

Nouns in Ega consist of a prefix and a nominal base, which can be basic, derived, or compound. Bole-Richard (1982, 379) also uses *qualified* as a type of nominal base. His example (Bole-Richard 1982, 390) for a qualified noun, however, does not structurally resemble the qualified noun phrase *these white cows* given above because there the adjective follows the noun whereas here the adjective is inserted between the noun and the prefix. Therefore, it is more likely to be a compound consisting of an adjective and a noun:

Table 19: Potential Ega compound noun

Ega	U-	g<a´g<a´	wla´	ma´
Constituent	PREF	qualifying adjective	qualified nominal root	DEM&PROX
Gloss	PREF	big	hen	DEM&PROX
English	this big hen			

Noun prefixes occur in various forms:

- as the morphophonemes I, E, O, U, and A
- as zero prefix
- as other rare prefixes

The table below gives examples and glosses for the prefixes.

Table 20: Ega noun prefixes

Prefix	Example	Gloss
i	i`gbe-	door
I	I`gbO-	hole
e	e-fi´	ear
E	E-na-	animal
o	o-te-	tree
O	O`tO`	father
u	u`du´	debt
U	U`g<lO`	village
a	a-nu-	two heads
zero prefix	J\i-Je`	sea
other	[no example]	[no example]

According to Bole-Richard, nouns with zero prefix are words which come from languages other than Ega, e.g. European languages or other African languages.

The category “other rare prefixes” is proposed by Gibbon/Ahoua/Salffner and unfortunately is not illustrated with an example.

For the prefix *a*, Bole-Richard only gives examples in plural. According to him, *a* does not occur in singular as a prefix. There are, however, mass nouns like *a-d<u´* meaning *water* or *a`NO`* meaning *air* which begin with *a*. None of the sources state whether these initial consonants are morphemes or part of the root. If they are morphemes, they may be part of a noun class characterised by *a* as a prefix. A hint towards *a* being a prefix is the word *e-d<u´*, which like *a-d<u`* means *water* but is semantically slightly different – *a-d<u`* means *water to drink* and *e-d<u´* means *water from rivers or rain*. Both lexemes may have the same root and just differ in their prefixes.

Prefixes pose constraints on the declension pattern for number and define the class of a noun. Class itself, however, is not that relevant for morphology, it only becomes relevant for syntagmatic relations:

- concord within nominal phrases
- concord between nominal phrases and predicative
- concord with pronominal anaphora

Ega uses two numera – singular and plural. Change in number is indicated by a change in the prefix. Possible singular prefixes are I, E, O, U. Possible plural prefixes are a, I and E. Prefixes must change, there is dissimilarity constraint for declension in number. Altogether, Bole-Richard lists seven different declension patterns:

Table 21: Noun prefixes – declension patterns

Singular prefix	Plural prefix
I	a
E	a
E	I
O	a
O	I
O	E
U	I
U	E

According to Bole-Richard, the noun may have a suffix, too, which functions as a determiner. Gibbon/Ahoua/Salffner, however, classify these morphemes as an independent part of speech, namely a determiner. This makes the morpheme part of the noun phrase, therefore it has been discussed above in the syntax section.

6.5. Ega phonetics and phonology

The Ega phonology has been described in more detail in a number of publications. (Bole-Richard 1982, Dago n.d., Bole-Richard n.d.[b], Connell/Ahoua/Gibbon 2002). For a brief outline on Ega phonetics and phonology cf. section 1.1.3.

6.6. Ega orthography

So far Ega is not a written language and there is no officially accepted orthographic system. Therefore it is not possible to describe the Ega orthography here.

There are, however, three competing proposals. Firstly, there is an orthography primer by Bole-Richard (n.d. [a]), which introduces Ega graphemes and offers reading exercises. Secondly, there is a system used by Gbery (n.d.) in his Ega dictionary. Finally, there are ILA (1996) proposals for an all-Ivorian system of writing. The various proposals are compared and analysed below in section 9.1.

7. STRUCTURE OF THE PRIMER

This sections is concerned with progression of the material for the learners. The questions to be discussed are

- Should there be a given progression? Why or why not?
- If yes, on which principles should progression be based for which fields?
- If no, which alternatives are there?

A linear progression for a primer is hard to define. If linguistic categories are used, there have to be good reasons for which categories to use and why to use them. Even if these categories are given, there may still not be an answer for endangered languages when the documentation is not complete. A usage-based approach, on the other hand, can already be used at a stage where it has been established in which situations, with which topics and for what purpose the language is used. It defines what is important to teach but still does not say anything about which order is best. What is more, there are the diverging needs of the various target groups which means that different orders of the content will be needed for the different learner groups.

As a consequence, I suggest a primer which consists of a number of modules which are organised like a lexicon. A lexicon is not usually read in a linear way. Instead, the user goes straight to the entry s/he needs. For navigation within the lexicon, the reader makes use of the lexicon's macro-, meso- and microstructure. These concepts can be applied to a primer as shown in Table 22.

Table 22: Equivalences between lexicon and primer constituents

Lexicon Constituents	Primer Constituents
lexical entry	module
microstructure	order within a module
mesostructure	references between modules
macrostructure	overall order of the primer modules

The advantage of the modular structure is that one set of modules can be used for all learners despite their different needs. It is easily extendable without having to worry about a prescribed structure. It is neither fixed nor prescriptive and instead it is flexible, easily extendable and can be adapted to the learners' diverse needs.

7.1. Microstructure

A unit in the primer can be introduced by a dialogue or story illustrating a functional usage of the language. If a language is used only in the spoken modality, a spoken dialogue is an appropriate starting point. Progression then automatically follows from this dialogue – unknown structures which occur in the dialogue have to be taught next.

A functional usage module may consist of only the dialogue or of the dialogue plus exercises on vocabulary, grammar etc. which occur in this particular dialogue. Specially made exercises which build on a particular dialogue may make learning easier, separate exercises keep the primer more flexible.

Possible topics for exercises are:

- Orthography: tones, di- and trigraphs vs. sequences of the respective monographs
- Listening: minimal pairs (tonal, phonemic) to train discrimination, letter-sound mapping, e.g. minimal pairs in listening and writing and how the differences are

represented in the writing

- Pronunciation: minimal pairs, implosives
- Grammar: verbal inflexion, interrogative sentences
- Vocabulary: semantic fields
- Pragmatics: routine formulae, formal and informal register
- Cultural background: riddles, proverbs, stories

7.2. Macrostructure

Macrostructure describes which entries or modules the primer includes and in which order they occur. I suggest reference modules and content modules. The reference modules contain a table of contents, an index and a preface which can be understood as a manual giving information on the content, structure, layout and aims of the primer and which explains how to use the primer.

Content modules can either address one aspect in one field of linguistics or they can be module clusters which deal with one topic in various fields of linguistics. These clusters can be

- dialogues
- phonetic/phonological issues occurring in the dialogues
- morphosyntactic issues occurring in the dialogues
- vocabulary occurring in the dialogues
- pragmatic issues occurring in the dialogues
- cultural background inherent in the dialogues

The order in which the modules are used depends on the learners, their interests and their needs. There is no fixed order of modules.

7.3. Mesostructure

The mesostructure describes how the modules are related to each other and to external sources. In the modular system, a comprehensive mesostructure is very important as it is a primary means of navigation.

References within the primer can link any module of the primer with any other module. Thus, there can be links between dialogues, exercises and explanations. Also, there can be references to parts of modules, e.g. not to a complete module but a particular exercise within a module.

8. LAYOUT OF THE PRIMER

This section answers the question:

- How is the teaching material presented to the learners?
- Which modalities are used in the presentation?
- Which media are used in the presentation?

The presentation format of the primer differs for every target group because the intended users

- differ in prior knowledge
- vary with respect to external factors (time available, motivation, chances of using Ega outside class etc.)
- have different needs
- are equipped with different infrastructure

8.1. Potential modalities

Gibbon/Mertins/Moore (2000, 438) define *modality* as the way a communicating agent conveys information to a communication partner. Applied to the primer *modality* describes the way the primer is presented to the learners.

There is a range of possible modalities for the primer. Firstly, there is the oral modality for learners who just want to understand and speak the language. Secondly, there is a written modality for those who also want to read and write in the language. Thirdly, there is a "real-life modality" for example in cases where the context of the teaching is the village storyteller's house. The real-life modality combines oral, visual, gestural and other possible modalities. It is particularly important for passing on ethnic identity and traditions. Fourthly, there is a virtual reality modality which uses computers and hypertexts. This modality combines oral, written, visual, auditory and other possible modalities and is suitable for future learners when the language is extinct. Computers and hypermedia can preserve the documented and described data of the language and one day present it to interested language revitalisers.

8.2. Potential media

Media are understood by Gibbon/Mertins/Moore (2000, 436) as the physical devices which are used to transmit or present information in communication.

Potential media for the primer are deduced from the potential modalities given above and from the infrastructural constraints.

8.2.1. Radio course

For the oral modality, the radio is suggested as the broadcasting medium. This has the advantage that literacy does not have to be taught first or simultaneously. In addition, the language is taught in the modality it is actually used in (although not in the context it is actually used in now). The radio is also a cheap means of reaching people. Provided that there are radios and there is electricity or that there are wind-up or solar radios, transmissions can be easily and widely received. For Ega, this is interesting because the villages which do have electricity and thus potentially also radio are exactly the villages where Ega is less widely spoken and more endangered than in the centre of the region.
Finally, radio courses have already been designed for other languages, e.g. for Welsh by the BBC (Jones/Davies/Davies 1995) or for German by the Deutsche Welle (1997) so that there is some expertise which can be drawn on.

8.2.2. Paperback book

For a combined oral, written and real-life modality, a module of paperback book units taught by teachers is suitable. This medium is backed up by using black boards, supplementary material, e.g. for literacy teaching, real-life artefacts or other material. A paperback book has the advantage that it is rather cheap, easy to publish and to distribute and that it can be used by many people. Also, it is a medium which is well known to learners and well established for teaching so that both teachers and learners can be expected to know how to use it.

8.2.3. Hypertext

For the virtual modality, computers and hypertexts are suitable. A hypertext is an ideal implementation of the modular structure of the primer. *Hypertext* is a term for a collection of documents which contain cross-references or "links" which allow the reader to move easily from one document to another. (Howe 2003) This non-linear structure and non-linear navigation is exactly what is intended with the modular structure.
As storage media, CD-ROMs and DVDs are well suited. They are not easily damaged by the climate conditions of for example tropical Africa, they can store a lot of data and can be used without access to the internet. CD-ROMs and DVDs can be used in the

permanent absence of native speakers, which makes them an ideal medium for distance learning. This is good for linguists who are interested in the language but cannot go there, i.e. are distant in space, and for future learners who will be distant in time.

At the moment, however, a hypertext is not a very useful medium for teaching Ega because the Ega do not have enough access to computers. Also, it could be argued that they do not have the expertise to use hypertexts. Still, I consider this less of a problem because expertise will develop as soon as the Ega start using this new medium.

9. APPLICATION TO EGA TEACHING MATERIALS

This chapter aims at putting two of the issues discussed above into practice. Firstly, it applies the ideas for orthography design to the proposals for an Ega orthography in order to decide on a well-founded theory-based, evaluated and justified proposal.

Secondly, the ideas on content, layout and module-internal structure are implemented in order to design a model unit for an Ega primer.

9.1. An Ega orthography

There are already three orthographic writing systems in use for Ega. Therefore this section will follow the steps suggested by Coulmas (1999) but instead of drafting a fourth proposal, this section compares the existing proposals and examines them by using the criteria set up in section 5.6.3.4. On the basis of this examination, I will decide which writing system or combination of systems will be used for the primer.

There are no publications explicitly addressing the question of an Ega writing system but there are two Ega specific orthographies and there are all-Ivorian guidelines.

ILA (1996) has set up an alphabet with sound-letter correspondences and orthographic rules which they suggest to use for all the Ivorian languages.

Gbery (n.d.) uses a system of orthography in his Ega dictionary in the orthography column. He does not explicitly deal with any orthographic problems. Whatever is mentioned here has been inferred from his usage of the orthography in his lexicon.

Bole-Richard (n.d. [a]) outlines a system of writing Ega in his primer “Je lis et j´écris l´éga”. He directs his instructions at an audience who is interested in learning how to read and write and not at the scientific community. Therefore, any meta description given here is also deduced from his orthography primer.

9.1.1. Step 1 – Phonological analysis

The relevant phonological categories in Ega on the segmental level are phonemes and tonemes. An overview of the Ega phonological system has been sketched in section 1.1.3.1.

9.1.2. Steps 2 and 3 – Proposals for mapping and augmenting

9.1.2.1. Writing system

All three proposals employ a phonemic writing system. Logographic elements are also included, namely punctuation (ILA 1996, 14) and capitalisation (ILA, Bole-Richard). The primer for another Ivorian language, Wan, also introduces Arabic numbers. (Mathias/Bolli 1996, pp. 47) This system can be taken for Ega, too.

9.1.2.2. Script and graphemes

All three proposals for Ega use a combination of the Roman alphabet and the International Phonetic Alphabet as their script. Using the Roman alphabet has the advantages that

- the Roman alphabet can be written on standard keyboards of ordinary typewriters. This reduces costs and enhances the chances for publications in the new language.
- the official language of the country, French, uses the Roman alphabet and therefore transfer can take place in both directions.
- other Ivorian languages use a combination of the Roman alphabet and the International Phonetic Alphabet, e.g. Agni (Anderson, Jérémie, Burmeister 1985) or Wan (Mathias/Bolli 1996).

The proposals differ with respect to how the characters are combined into graphemes and how the graphemes are mapped onto the phonemes.

Table 23 gives an overview of Ega vowels and their respective graphemes as suggested by Bole-Richard (n.d. [a]), Gbery (n.d.) and ILA (ILA 1996, pp. 6).

Table 23: Ega vowels and proposals for their grapheme counterparts[12]

Phoneme	Grapheme		
	Bole-Richard	Gbery	ILA
a	a	a	a
e	e	e	e
E	ɛ	ɛ	ɛ
i	i	i	i
I	ɪ	ɪ	ı
o	o	o	o
O	ɔ	ɔ	ɔ
u	u	u	u
U	ʊ	ʊ	ʊ

Except for /I/, all three proposals choose the same graphemes. The differences between the graphemes for /I/ are minimal.

Table 24 gives an overview of Ega consonants and their respective graphemes as suggested by Bole-Richard (n.d. [a]), Gbery (n.d.) and ILA (ILA 1996, pp. 6). Differences between the proposals are indicated by italics.

Table 24: Ega consonants and proposals for their grapheme counterparts

Phoneme	Grapheme		
	Bole-Richard	Gbery	ILA
p	p	p	p
t	t	t	t
c	c	c	c
k	k	k	k
kp	kp	kp	kp
b	b	b	b
d	d	d	d

12 The font used for the graphemes which have been taken from the International Phonetic alphabet is Ipa-samd Uclphon1 SILDoulosL.

Phoneme	Grapheme		
	Bole-Richard	Gbery	ILA
J	*j*	*ʄ*	*j*
g	g	g	g
gb	gb	gb	gb
b<	bh	bh	bh, ɓ
d<	dh	dh	dh, ɗ
J<	*jh*	*ʄh*	*[not indicated]*
g<	gh	gh	gh, ɣ
gb<	gbh	gbh	[not indicated]
f	f	f	f
s	s	s	s
x	x	x	x
v	v	v	v
z	z	z	z
l	l	l	l
j	*y*	*j*	*y*
w	w	w	w
m	m	m	m
n	n	n	n
J	*ny*	*ɲ*	*ny, ɲ*
N	*ng*	*ŋ*	*ng, ŋ*
r[13]	*[not indicated]*	*r*	*r*
voiced velar fricative[13]	*[not indicated]*	*ɣ*	*gh*

For the majority of the consonants, the graphemes in the three proposals correspond. Differences occur with the following sounds:

- palatal voiced plosive, implosive, approximant and nasal
- velar nasal

13 These sounds do not occur in the Ega phonetic inventory (Ahoua/Connell/Gibbon 2002, pp. 100) and not in Bole-Richard's primer either. Ahoua/Connell/Gibbon explicitly say that they consider these sounds to be allophones of [l] and [x] respectively. Bole-Richard apparently does so, too, but does not explicitly say that.

- velar voiced fricative
- alveolar lateral approximant

In addition, there are overlaps regarding the graphemes *j* and *ɣ*. Gbery uses the grapheme *j* for the palatal voiced approximant. Bole-Richard on the other hand uses *j* for the palatal voiced plosive. The grapheme *ɣ* is used by Gbery for the voiced velar fricative, whereas ILA suggests using it for the voiced velar implosive.

Interestingly, ILA suggests the IPA symbols as graphemes for the bilabial and dental implosive but not for the velar implosive, which is inconsistency within the ILA proposal. However, this is not the place to discuss the ILA system in general, what is relevant here is whether the system is suitable for Ega. Therefore this inconsistency is only mentioned and not explored in more detail.

The ILA proposal lacks two of the sounds needed to write Ega, namely the palatal and the labiovelar implosive. These phonemes are very rare (Connell/Ahoua/Gibbon 2002) and therefore ILA may not have been aware of the fact that these phonemes would occur and be needed.

Finally, Gbery uses additional letters (*r* and *ɣ*), which Bole-Richard does not include in his orthography primer. In the Ega phonemic inventory (Connell/Ahoua/Gibbon 2002), the graphemes *r* and *ɣ* do not have phoneme counterparts. Instead, they are allophones of /l/ and /x/ respectively.

Table 25 gives an overview of Ega tonemes and their respective graphic realisations as suggested by Bole-Richard (n.d. [a]), Gbery (n.d.) and ILA (ILA 1996, pp. 6).

Table 25: Ega tones and proposals for their grapheme counterparts

Toneme	Grapheme		
	Bole-Richard	Gbery	ILA
‘	ˊ	[not marked]	ˊ
-	[not marked]	[not marked]	[not marked]
ˋ	ˋ	[not marked]	ˋ

Gbery does not mark tone at all. Bole-Richard and ILA both use diacritics to augment the Roman alphabet so that it accommodates tonal information. Both Bole-Richard and ILA treat the middle tone as the default case and do not explicitly mark it. The high and low tone are marked by *accent aigu* and *accent grave* respectively.

9.1.2.3. Orthographic rules

Gbery's orthography does not contain any information on orthographic rules. It may be possible to analyse his usage of word segmentation e.g. for compounds or reflexive verbs and deduce rules from that but this would go beyond the needs and the time available for this thesis. Therefore Gbery is not included here.

Issues explicitly addressed are word segmentation, punctuation and capitalisation.

Regarding word segmentation, ILA (1996, 14) comments on elision and proposes marking elision with an apostrophe. In cases where words are divided at the end of a line, a hyphen or a plus are suggested. Bole-Richard (n.d. [a]) goes into more detail and puts forward rules on the correct spelling of

- verbs and their various possible complements
- negated verbs
- elisions which are not supposed to enter the written system
- elision which may enter the written system under certain conditions
- compound nouns vs. noun plus complement

Regarding punctuation, ILA (1996, 14)suggests using the French system punctuation in the same way as it is used in French. Bole-Richard does not comment on punctuation.

Regarding capitalisation, ILA (1996, 14) proposes using capital letters at the beginning of sentences and for proper names.

9.1.3. Evaluation

9.1.3.1. Social Acceptance

Social aspects which might have an effect on the orthography design have not been investigated. If religion is to play a role, choosing the Roman alphabet is justified as the Ega are mainly Catholic (SIL 2002) and the Christian church has been associated with the Roman alphabet by Coulmas (1996, 138). Politics and identity may be factors in the way that the Ivory Coast as a former French colony still has strong ties with France and uses French as its official language. Whether this is an argument for or against using a script similar to French has not been looked into but for the purpose of this thesis it will be seen as not being an obstacle.

Furthermore, no comments can be made on the practical social acceptance because there have not yet been any Ega literacy classes.

9.1.3.2. Precision and Economy

Regarding tonemes, neither of the proposals explicitly marks all tonemes. The ILA and Bole-Richard proposals however guarantee precision in so far as all three tones can be identified unambiguously – the default middle tone by zero marking and the high and low tone by accent marks.

Gbery does not indicate tones at all. Why he leaves them out cannot be said. A possible line of argumentation is that tone need not be marked because the meaning of a word can be inferred from the context without explicitly marking lexical or grammatical tone. This is similar to word stress which distinguishes words for example in German or English and is not marked orthographically. Whereas speakers with a good command of the language are able to do infer the right pronunciation from the context, this causes difficulties for language reacquirers and may be difficult for language maintainers. These learners learn the actual language and its system of writing simultaneously. Thus, they are likely to have problems with inferring the meaning from the context and in placing the right tones on the syllables by themselves because they may not know the lexical and grammatical tone yet and have to learn it as they go along.

9.1.3.3. Phonological accuracy

None of the proposals follow the principle of “one letter – one sound”. Instead of using only monographic letters, the suggestions include also digraphs and trigraphs which consist of letters that sometimes also occur as monographs, e.g. the trigraph *gbh* vs. the monographs *g* and *b*.

If the word *sound* in the principle “one letter – one sound” is understood as meaning *phoneme*, then Gbery’s proposal does not follow this principle because he distinguishes between allophones so that there are two letters for one sound. The ILA principles are ambiguous in a similar way because they suggest alternative spellings for one phoneme – for the phoneme /d</ they suggest the graphemes *dh* and *ɗ*. In addition, they propose an ambiguous grapheme representing two different sounds, i.e. *gh* standing both for a velar voiced implosive and for a velar voiced fricative.

The letter *h* is used consistently for marking implosives by Gbery and Bole-Richard. This is helpful because it contributes to transparency and it is phonologically consistent because it keeps the place and voicedness of the plosive counterpart and just adds the implosive marker.

Taken together, the orthography proposal by Bole-Richard is the only unambiguous proposal. It is phonologically accurate along the (adapted) principle of "one grapheme – one phoneme".

9.1.3.4. Transferability

French and other Ivorian languages use a script which is very similar to the proposals for Ega. Therefore positive transfer can take place, whether it actually does is a question which can only be resolved in an empiric study.

9.1.3.5. ASCII Characters

For [–ATR] vowels, all proposals make use of phonetic symbols rather than ASCII characters. For consonants, Bole-Richard uses only ASCII characters, the ILA proposals can also be applied in a way that the alphabet only contains ASCII characters. Gbery uses both phonetic and ASCII characters. For tonemes, all the diacritics needed are included in the ASCII set.

As the Ega vowel inventory contains nine different vowels, the ASCII set, however, only offers five vowel characters, it has to be augmented. One solution is drawing on SAMPA characters and using capital letters for [–ATR] vowels. Then there would be uniquely identifiable phoneme-grapheme-matches but in return it would cause problems with capitalisation. There is bound to be a number of other ways out but there is always the problem that using simple characters is not the only criterion and the other principles have to be obeyed too.

Also, as many Ivorian languages use phonetic symbols for vowels, the publishers in the Ivory Coast may have the fonts available. For other ways of publishing, e.g. via HTML on the internet or in print media marked-up with LaTeX, phonetic symbols can be added when the necessary package is installed or the appropriate backslash command is used.

9.1.3.6. Learnability and practicality

As the Ega orthography has not yet been taught and not yet been used by people other than linguists, nothing can be said about its learnability or practicality. This can be investigated once the system of writing has been learned and is used in real life and once there are criteria for *good* or *bad learnability* and for *highly* or *hardly practical*.

9.1.4. Making a choice

Regarding vowels, all three proposals agree and I do not have any better suggestions regarding the phonetic symbols. Therefore the vowels are left the way they are.

Regarding consonants, Gbery and ILA use graphemes which do not occur in the ASCII set. Furthermore, the ILA suggestion lacks graphemes for two phonemes and is inconsistent with other graphemes. Bole-Richard's proposal are unambiguous and use ASCII characters only, therefore they will be used as the graphemes for the primer.

Regarding tones, the primer will take up the ILA and Bole-Richard suggestion to explicitly mark the high and low tone and to set the middle tone as default. This is considered best practice to suit the needs of learners who have to (re)learn both the language and its writing system simultaneously.

Regarding punctuation and capitalisation, the primer follows the ILA guidelines. They are the only available proposals, they use ASCII characters and are used in the same way as in French, which contributes to transferability.

In terms of word segmentation, the only comprehensive system of rules available is Bole-Richard's. Still, it is difficult to make a decision on this issue because this requires complex morphosyntactic analysis and description before it can be implemented and the linguistic basis available so far is sufficient for this.

9.2. A unit in an Ega textbook

9.2.1. Greetings

The communicative situation chosen for the primer is greetings. This is because greetings are universal and essential communicative situations which occur in every culture. (Eibl-Eibesfeld/Senft 1987, 96)

The greeting module is needed by beginners, i.e. mainly by language reacquirers. In addition it can be used by language maintainers if they have difficulties in this communicative situation. Language developers may find it useful because greeting procedures are very basic and therefore well-known and can be used to raise meta-linguistic awareness of the language without having to learn the language itself simultaneously.

What is needed for greetings is

- for usage: the functional use of the language for the discourse type greeting, i.e.

the semantic frame, semantic script and the lexical material

- for vocabulary: possibly the times of the day
- for phonology: phonologically relevant features which occur in the phrases (e.g. implosive alveolar, tones, possibly the palatal nasal)
- for morphology: formality marker, plural marker
- for syntax: declaratives

The introduction module consists of a collection of matching pictures and dialogues. The dialogues are constructed so that they include examples of all the variables time of the day, number and register. The pictures are drawn in a way that they reflect and illustrate the meaning and characteristics of the dialogue.

For times of the day, I use the rising sun, the position of the sun at midday and the moon and stars respectively. For number, I use a single person vs. two people or a group of people. I have no data which shows what to use for register. Therefore I will rely on the German concept for the time being and use a child and an adult. I am aware of the fact that this may be wrong and may have to be changed. An example of the pictures suggested for the primer is given below.

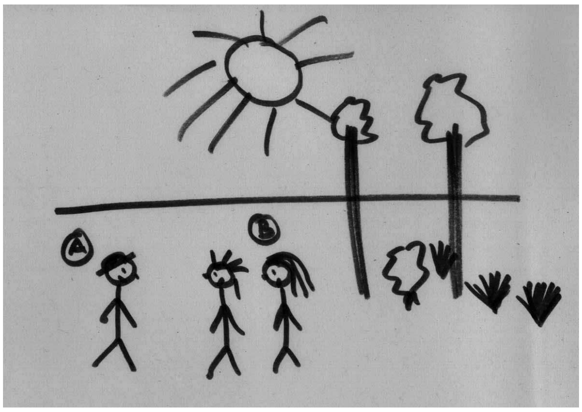


Figure 8: A picture for the primer

The syntax module could consist of the same dialogues and pictures but here the learners are asked to find out what is similar and what is different about the dialogues. They can then mark the “anchor” *mU´ka*` in one colour and the other constituents in other colours. This is meant as a basis for the learners for working out the motivations behind the various constituents and for explaining the meaning and the structure of the phrases. In the end, the learners are expected to be able to construct a greeting template which they can use for reference.

For practicing the greeting formulae I suggest exercises for

- filling gaps for individual constituents
- giving replies to a greeting formula or giving the original formula to a reply (with the help of pictures to illustrate what is intended)
- role-plays (or why not go on one-day excursions to another Ega village and greet everybody at different times of the day?)

9.2.2. Story telling

Teaching story telling involves teaching the language and transmitting ethnic history, traditions of the community and ethnic identity. These are high aims and include more than either traditional first language teaching or second language teaching. Then again, transmitting cultural values and traditions is something which has been done by the community itself for a long time already so that there is expertise in transmitting traditions and values which can be used.

Learning how to tell stories is suitable for language maintainers and language developers. It is less interesting for language reacquirers in their beginning stages as they lack the command of the language to be able to understand or tell a story.

The Ega have their own way of teaching story telling. Children learn by observing, copying and doing. In addition to that, the village storyteller may choose to explicitly teach children a particular story. In this case, the storyteller said that he would call two boys to him every night and he would tell them a story. Afterwards they tell him either the same story or another one and the storyteller listen to them. If they forget something, the storyteller corrects them and they practise until they do not forget anything anymore. Learning therefore means listening, repeating and being corrected and the storyteller's function is to show the children how to tell the story better. (Rafferty/Salffner 2002, 39)

It therefore seems sensible to teach story telling in co-operation with the village storyteller as he knows best and has his ways of instructing the youth. In addition, this is an excellent opportunity to include the community in the (institutional) maintenance efforts. It integrates traditional and institutional teaching and thus fits in the paradigm for primer making and teaching outlined in section 5.7.

10. Conclusion and Outlook

10.1. Summary of results

Primers are closely related to language documentation and must be integrated in an overall model of language documentation so that both language documentation and teaching can benefit from developments and advancements in the other discipline. Issues discussed in language documentation and quality criteria and parameters for the selection of communicative events used in language documentation also apply to primers.

The textlinguistic document model which includes the 3+1 components content (semantics), structure (syntax), rendering (pragmatics) and constraints on the document proved to be applicable and useful as an organisational matrix. It is general enough to cover a range of languages and at the same time it can be adapted adequately to specific languages.

The content of a primer for a spoken endangered language is to be based on pragmatic or functional usage of the language. Relevant structures from the other levels of linguistic analysis follow from this functional usage – whatever structures occur in for example a greeting dialogue become a topic in the primer. This empiric, functional and pragmatic approach is to be combined with a general theoretical and formal approach which serves as a tool for guaranteeing comprehensiveness and as a basis for work when there is not enough empirical data.

Regarding the structure or progression of a primer, a linear progression does not meet the constraints of the mixed learner group in a community speaking an endangered minority language. Instead, a modular structure is more suitable where the individual components are linked with each other and where learners and teachers can pick exactly those elements which are useful for them. Similarly, the ways of teaching and the material used for the teaching have to match the interests, objectives and needs of the respective learner group. A predetermined medium is equally inappropriate as a fixed structure.

10.2. Problems and open issues

Because of lack of data and the need for limitation, the discussion of the primer structure and layout had to be kept short. The practical application to actual Ega teaching material could not be implemented for the same reasons.

It cannot be guaranteed that the proposals for an Ega orthography for the primer are useful because they have been neither used nor evaluated. Also, it cannot be guaranteed that the linguistic material used for the Ega greeting module is correct as it has not been verified by a native speaker.

In a next phase, more modules have to be set up and the material has to be implemented in actual Ega language classes so that it can be evaluated and modified if necessary. Currently this is not possible because of the unstable political situation in the Ivory Coast.

10.3. Outline of future avenues of work

The frame of reference established in this thesis has to be taken to appropriate platforms where it can be discussed by other linguists and users so that the model can be extended and modified where there is need for improvement. Also, elements regarding the teaching methodology, a teacher perspective and the practical application and organisation could be added.

What is most important, however, is that the need for a well-founded and holistic linguistic basis for modelling primers for endangered spoken minority languages is recognised among linguists, primer designers and teachers and that appropriate environments such as journals, organisations or courses of studies are established where these issues are investigated. There are promising outsets such as an M.A. course on language documentation and description offered by the School of Oriental and African Studies at London University (2003, 4) which includes for example language curriculum design as an element of one of its core courses. Hopefully programmes like this will raise awareness of the needs of endangered languages and hopefully some people some day are prepared to come up with the money to pay for this kind of language teaching so that some day every person can get what is their human right – education in one's own language.

11. BIBLIOGRAPHY

Agatucci, Cora, (n.d.). *African Storytelling: Oral Traditions*. URL: http://www.cocc.edu/cagatucci/classes/hum211/afrstory.htm, accessed in September 2001.

Ahoua, Firmin (2003). *Nominal Morphology of Ega*. Technical Report. Bielefeld: Universität Bielefeld.

Ahoua, Firmin; Connell, Bruce; and Gibbon, Dafydd (2000). *Ega. A documentation model for an endangered Ivorian language. Project description, methods, and scheduling*. Workshop Presentation. Nijmegen: MPI.

Ahoua, Firmin; Connell, Bruce; and Gibbon, Dafydd (2001). *Ega. A documentation model for an endangered Ivorian language*. DOBES Pilot Phase Workshop Presentation. Bielefeld: Universität Bielefeld.

Ahoua, Firmin; Connell, Bruce; and Gibbon, Dafydd (2002). *Multilingual factors in endangerment space. Language shift in the Ega community*. Presentation at the DGfS Jahrestagung. Mannheim.

Aikhenvald, Alexandra Y. (2003). Teaching Tariana, an endangered language from northwest Amazonia. *International Journal of the Sociology of Language* 161, 125-139.

Anderson, Anoh Kouao; Jérémie, Bouadi Kouamé; and Burmeister, Jonathan (1985). *Syllabaire en Agni Sanvi*. Abidjan: Institut de Linguistique Appliquée and Societé Internationale de Linguistique.

Bamgbose, Ayo (1984). Minority languages and literacy. In *Linguistic Minorities and Literacy. Language Policy Issues in Developing Countries*, Florian Coulmas (ed.), 21-28. Trends in Linguistics. Studies and Monographs 26. Berlin, New York, Amsterdam: Mouton Publishers.

Blass, Armin and Wolf Friedrich (1969). *Englischer Wortschatz in Sachgruppen*. München: Max Hueber Verlag.

Bole-Richard, Rémy (1982). Esquisse de grammaire: L´Ega. In *Atlas des Langues Kwa. Tome 2*, Georges Hérault (ed.), 359-401. Abigjan: Institut de Linguistique Appliquée.

Bole-Richard, Rémy (n.d. [a]). *Je lis et j´écris l´éga*. Unpublished orthography primer.

Bole-Richard, Rémy (n.d.[b]). *Une autre approche de l´harmonie vocalique. Le mot phonologique*. Unpublished sketch. URL: http://www.spectrum.uni-bielefeld.de/langdoc/EGA/EgaCatalogue/ega_data_2001.html, accessed in

March 2004.

Bole-Richard, Rémy (n.d.[c]). *Lexique Ega*. Unpublished Ega Lexicon. URL: http://www.spectrum.uni-bielefeld.de/langdoc/EGA/EgaCatalogue/ega_data_2001.html, accessed in March 2004.

Bubenik, Vit (1999). *An Introduction to the Study of Morphology*. LINCOM Coursebooks in Linguistics 07. München, Newcastle: LINCOM Europa.

Collins Cobuild English Dictionary (1995). London: HarperCollins Publishers.

Connell, Bruce; Ahoua, Firmin; and Gibbon, Dafydd (in press). Ega. A preliminary assessment of endangerment. In *Trends in African Linguistics*, L.M. Hyman and Ian Maddieson (eds.). Trenton NJ: Africa World Press.

Connell, Bruce; Ahoua, Firmin; and Gibbon, Dafydd (2002). Ega. Illustrations of the IPA. *Journal of the International Phonetic Association* 32, 99-104.

Coulmas, Florian (1981). *Routine im Gespräch. Zur pragmatischen Fundierung der Idiomatik*. Wiesbaden: Akademische Verlagsgesellschaft Athenaion.

Coulmas, Florian (1984). Linguistic minorities and literacy. In *Linguistic Minorities and Literacy. Language Policy Issues in Developing Countries*, Florian Coulmas (ed.), 5-20. Trends in Linguistics. Studies and Monographs 26. Berlin, New York, Amsterdam: Mouton Publishers.

Coulmas, Florian (1999). Development of Orthographies. In *Literacy. An International Handbook*, Daniel A. Wagner, Richard L. Venezky, and Brian V. Street (eds.), 137-142. Colorado, Oxford: Westview Press.

Coulmas, Florian (2003). *Writing systems. An Introduction to their Linguistic Analysis.* Cambridge: Cambridge University Press.

Council of Europe (2001). *A Common European Framework of reference for Languages. Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Crystal, David (1995). *Die Cambridge Enzyklopädie der Sprache*. Translated and edited by Stefan Röhrich. Frankfurt/Main, New York: Campus Verlag.

Dago, Georgette (n.d.). *Phonetique Generale et physiologique*. Unpublished M.A. Thesis. http://coral.lili.uni-bielefeld.de/langdoc/EGA/, accessed in March 2004.

Deutsche Welle (1997). *Marktplatz*. Radiokurs. Köln: Deutsche Welle.

DOBES Ega Project. Nijmegen: MPI and Bielefeld: Universität Bielefeld. URL: http://www.mpi.nl/DOBES/teams/Ega/Ega.html, accessed in March 2004.

Edmondson, Willis; and House, Juliane (2000). *Einführung in die Sprachlehrforschung*. 2nd revised edition. Tübingen, Basel: Francke.

Ehnert, Rolf; and Hornung, Wolfgang (1986). Wie sollen Studierende des Faches Deutsch als Fremdsprache auf ihre künftige Tätigkeit als Fachsprachenlehrer vorbereitet werden? Entwurf eines Teilcurriculums. *Jahrbuch Deutsch als Fremdsprache* 12, 204-216.

Eibl-Eibesfeld, Irenäus and Gunter Senft (1987). *Studienbrief Rituelle Kommunikation: Kommunikation – Wissen – Kultur*. Hagen: Fernuniversität – Gesamthochschule Hagen, Fachbereich Erziehungs- und Sozialwissenschaften.

Feldmeier, Alexis (2003). Buchstabenprogression in der Alphabetisierung ausländischer Erwachsener. *Deutsch als Zweitsprache* 1, 26-32.

Flores Farfán, José A. (2001). Culture and Language Revitalisation, Maintenance, and Development in Mexico. The Nahua Alto Balsas Communities. *International Journal of the Sociology of Language* 152, 185-197.

Fromkin, Victoria; and Rodman, Robert (1998). *An Introduction to Language*. 6th edition. Orlando: Harcourt Brace College Publishers.

Gbery, Eddie A. (n.d.) Unpublished Ega lexicon. URL: http://www.spectrum.uni-bielefeld.de/langdoc/EGA/EgaCatalogue/ega_data_2001.html, accessed in March 2004.

Gibbon, Dafydd (1999). *Ega. Documentation model for an endangered Ivorian language*. Project Proposal. URL: http://www.spectrum.uni-bielefeld.de/langdoc/EGA/Proposals/Ega-proposal/index.html, accessed in March 2004.

Gibbon, Dafydd (2001). *The efficient documentation of endangered languages*. Presentation at the 33rd Poznan Linguistic Meeting. URL: http://www.spectrum.uni-bielefeld.de/langdoc/EGA/Presentations/EndangeredLangs.zip, accessed in March 2004.

Gibbon, Dafydd (2002). Ubiquitous multilingual corpus management in computational fieldwork. *Proceedings LREC 2002 Satellite Workshop on Portability.*

Gibbon, Dafydd (2003[a]). *Ega Morphosyntax*. Technical report. Bielefeld: Universität Bielefeld.

Gibbon, Dafydd (2003[b]). *Types of dictionary and their classification*. Class material. URL: http://coral.lili.uni-bielefeld.de/Classes/Summer2003/Lexicography/lexicography/node10.html, accessed in March 2004.

Gibbon, Dafydd (n.d.). *Ega SAMPA*. Technical Report. Bielefeld: Universität Bielefeld.

URL: http://coral.lili.uni-bielefeld.de/langdoc/EGA/Formats/egasampa.txt, accessed in March 2004.

Gibbon, Dafydd; Ahoua, Firmin; and Salffner, Sophie (2003). *Outline of Ega Morphosyntax*. Technical report. Bielefeld: Universität Bielefeld. URL: www.spectrum.uni-bielefeld.de/langdoc/EGA/, accessed in March 2004.

Gibbon, Dafydd; and Bows, Cathy (2003). *GOLD POS categories and their relevance to Ega mark-up*. Technical Report. Bielefeld: Universität Bielefeld.

Gibbon, Dafydd; Connell, Bruce; Ahoua, Firmin; and Urua, Eno-Abasi (2001). *The EGA documentation model. Evaluation and application to selected West African Languages*. Project Proposal. URL: http://www.spectrum.uni-bielefeld.de/langdoc/EGA/Proposals/Ega-proposal2/index.html, accessed in March 2004.

Gibbon, Dafydd; Mertins, Inge; and Moore, Roger K. (eds.) (2000). *Handbook of Multimodal and Spoken Dialogue Systems. Resources, Terminology and Product Evaluation*. Boston, Dordrecht, London: Kluwer Academin Publishers.

Häublein, Gernot; and Jenkins, Recs (1998). *Thematischer Grund- und Aufbauwortschatz Englisch*. Neubearbeitung. Stuttgart, Düsseldorf, Leipzig: Ernst Klett Verlag.

Himmelmann, Nikolaus P. (1993). Material Ambitions. *LW* 7, 66-68.

Himmelmann, Nikolaus P. (1998). Documentary and Descriptive Linguistics. *Linguistics* 36, 161-195.

Hopkins, Andy (2002). Users' guide for textbook and material writers. In *Common European Framework of Reference For Languages: Learning, teaching, assessment. A Guide for Users*, Sophie Bailly, Sean Devitt, Marie-José Gremmo et al. (eds.), 201-232. Strasbourg: Language Policy Division. URL: http://culture2.coe.int/portfolio/documents/Guide-for-Users-April02.doc, accessed in January 2004.

House, Juliane (1996). *Untersuchungen zu Unterschieden im Diskursverhalten englischer und deutscher Muttersprachler*. URL: http://www.tu-dresden.de/sulifg/daf/landesku/kapit3.htm, accessed in June 2003.

Howe, Denis (2003). *The Free On-line Dictionary of Computing*. URL: http://wombat.doc.ic.ac.uk/foldoc/foldoc.cgi?hypertext, accessed in March 2004.

ILA (1996). *Une Orthographe Pratique Des Langues Ivoiriennes*. Abidjan: Institut de Linguistique Appliquée.

Jones, Ann M.; Davies, Basil; and Davies, Cennard (1995). *The ABC of Welsh*. Radio Course Welsh. BBC Radio Wales.

Knowles-Borishade, A.F. (1991). Paradigm for Classical African Orature: Instrument for a Scientific Revolution? *Journal of Black Studies* 21, 488-500.

Kropp Dabuku, Mary Esther (1980). West African Language Data Sheets. Unpublished questionnaire.

Kühn, Peter (1979). *Der Grundwortschatz. Bestimmung und Systematisierung*. Reihe Germanistische Linguistik 17. Tübingen: Max Niemeyer Verlag.

Kuhn, Franz-Josef (1986). CA, CE, CI, CO, CU. Alphabetisierung in Ecuador. In *ABC und Schriftsprache: Rätsel für Kinder, Lehrer und Forscher*, Hans Brügelmann (ed.), 216-228. Reihe Libelle: Wissenschaft: Lesen und Schreiben. vol. 1. Konstanz: Faude Verlag.

Ladefoged, Peter (2003). *Phonetic Data Analysis. An Introduction to Fieldwork and Instrumental Techniques*. Oxford: Blackwell Publishing.

Langenscheidts Grundwortschatz Englisch. Ein nach Sachgebieten geordnetes Lernwörterbuch mit Satzbeispielen. (1987) Berlin, München, Wien u.a.: Langenscheidt.

Lehmann, Christian (1982). Directions for Interlinear Morphemic Translation. *Folia Linguistica* 16, 199-224.

Lehmann, Christian (2001). *Language Documentation. A program*. Unpublished paper. URL: http://www.uni-erfurt.de/sprachwissenschaft/personal/lehmann/CL_Publ/Language_documentation.pdf, accessed in March 2004.

Lehrer, Adrienne (1974). *Semantic Fields and Lexical Structure*. North-Holland Linguistics Series 11. Amsterdam, London: North-Holland Publishing Company.

Litvinov, Pavel P. (1999) *3000 anglijskij slov. Technika zapominanija. Tematičeskij slovar'-minimum*. Moscow: Rol'f: Ajris Press.

Lucien, Bazé; and Salffner, Sophie (2002). *Survival Ega*. URL: http://www.spectrum.uni-bielefeld.de/phonlab/survega.htm, accessed in March 2004.

Lüger, Heinz-Helmut (1999). *Routinen und Rituale in der Alltagskommunikation*. Fernstudienprojekt zur Fort- und Weiterbildung im Bereich Germanistik und Deutsch als Fremdsprache. Fernstudieneinheit 6. Berlin, München, Wien u.a.: Langenscheidt.

Mabe, Jacob E. (ed.) (2001). *Das Afrika-Lexikon. Ein Kontinent in 1000 Stichwörtern*. Stuttgart, Weimar: Verlag J.B. Metzler and Wuppertal: Peter Hammer Verlag.

Mathias, Bomisso Gbayoro; and Bolli, Margrit (1996). *Syllabaire Wan*. Abidjan:

Societé Internationale de Linguistique.

Mutere (n.d.) *Motif #1: Call and Response*. URL: http://artsedge.kennedy-center.org/aoi/resources/hg/motif1.html, accessed in January 2002.

Neuf-Münkel, Gabriele and Regine Roland (1994). *Fertigkeit Sprechen*. Erprobungsfassung 11/94. Fernstudienprojekt zur Fort- und Weiterbildung im Bereich Germanistik und Deutsch als Fremdsprache. Fernstudieneinheit. Berlin, München, Leipzig u.a.: Langenscheidt.

Nickolaus, Günther (1977). *Grund- und Aufbauwortschatz Französisch*. Stuttgart: Ernst Klett Verlag.

Obiechina, E. (1993). Narrative Proverbs in the African Novel. *Research in African Literatures* 24, 123-140.

Ouane, Adama (2002). In welcher Sprache soll/kann alphabetisiert werden, wenn die Muttersprache nicht die “Amtssprache” ist? In *Alphabetisierung und Sprachenlernen*, Thilo Fitzner (ed.), 71-54. Bad Boll: Evangelische Akademie und Stuttgart: Ernst Klett Verlag.

Payne, Thomas E. (1997). *Describing Morphosyntax. a guide for field linguists.* Cambridge: Cambridge University Press.

Pearsall, Judy (ed.) (1999). *The Concise Oxford Dictionary*. 10th ed. New York: Oxford University Press.

Poser, William J. (2002). Uses of Web Pages for Endangered Languages. *International Journal of the Sociology of Language* 158, 227-237.

Rafferty, Wayne; and Salffner, Sophie (2002). *Once upon a time in a country far, far away. Ritualisation and Ritualised Communication in African Orature*. Unpublished term paper written at Bielefeld University.

Reichmann, Oskar (1990) Das onomasiologische Wörterbuch. Ein Überblick. In *Dictionaries. An International Encyclopedia of Lexicography*. 1057-1067. Handbücher zur Sprach- und Kommunikationswissenschaft 5. Berlin, New York: Walter de Gruyter.

Salminen, Tapani (1999). *UNESCO Red Book On Endangered Languages. Europe*. URL: http://www.helsinki.fi/~tasalmin/europe_index.html, accessed in January 2004.

Scheub, Harold (1985). A Review of African Oral Traditions and Literature. *African Studies Review* 28, 1-72.

School of Oriental and African Studies/University of London (2003). *Course Handbook for MA in Language Documentation and Description and PhD in Field Lingusitics 2003-4. Endangered Languages Academic Programme*. Technical

Report. London: School of Oriental and African Studies, University of London.

SFB 268. (n.d.) *Kulturentwicklung und Sprachgeschichte im Naturraum Westafrikanische Savanne*. DFG Project. Frankfurt: Johann-Wolgang-Goethe Universität. URL: http://www.informatik.uni-frankfurt.de/~sfb268/, accessed in March 2004.

SIL (2002). Ega: a language of the Côte d'Ivoire. In *Ethnologue*. URL: http://www.ethnologue.com/show_language.asp?code=DIE, accessed in January 2002.

Srivastava, R. N. (1984). Linguistic minorities and national languages. In *Linguistic Minorities and Literacy. Language Policy Issues in Developing Countries*, Florian Coulmas (ed.), 39-46. Trends in Linguistics. Studies and Monographs 26. Berlin, New York, Amsterdam: Mouton Publishers.

Tallermann, Maggie (1998). *Understanding Syntax*. Understanding Language Series. London: Arnold.

Trask, Robert L. (1996). *A Dictionary of Phonetics and Phonology*. London: Routledge.

United Nations (1948). *Universal Declaration of Human Rights.* URL: http://www.un.org/Overview/rights.html, accessed in March 2004.

12. APPENDIX

The appendix consists of links to relevant documents and of a sample page for the Ega primer which introduces Ega greetings.

Gnama sociolingustic data (anonymised versions)

- Gnama village questionnaire
- Gnama children questionnaire

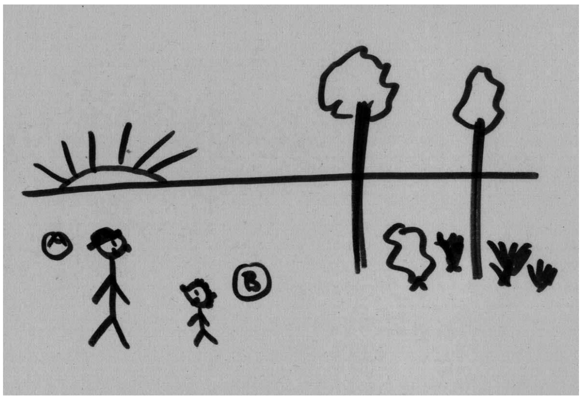
Ega story telling

- Video file in .avi format
- Annotation in .xml format

Ega core vocabulary

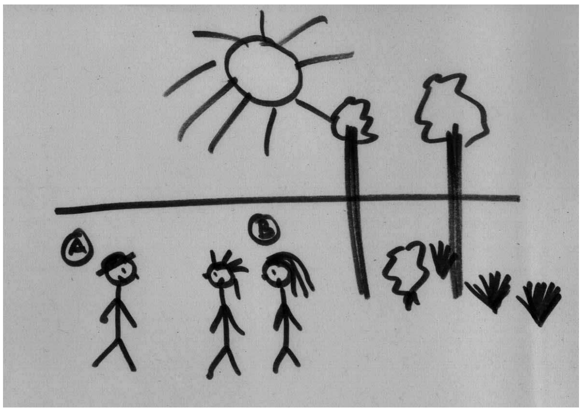
Proposal for an introduction to Ega greetings

Pour dire *Salut!* en Ega



A: ÓflI mÚkà!

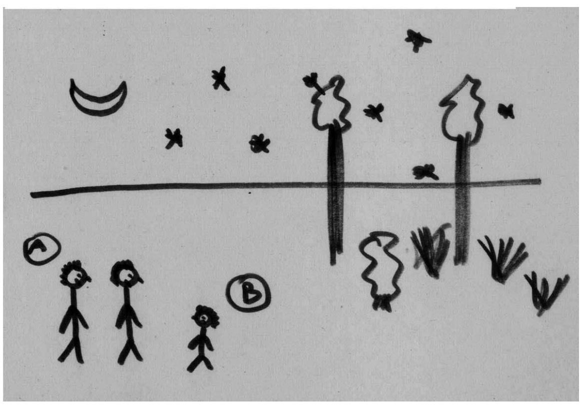
B: ÓflI mÚkà bE!



A: mÚkà nÚ!

B1: mÚkà!

B2: mÚkà!



A1: ónyodho mÚkà!

A2: ónyodho mÚkà!

B: ónyodho mÚkà nÚ bE!

ERKLÄRUNG

Ich versichere, dass ich die vorliegende Arbeit selbstständig angefertigt habe und keine Hilfsmittel und Quellen außer den von mir angegebenen verwendet habe. Alle Stellen, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen sind, habe ich unter Angabe der Quelle als Zitat oder sinngemäße Wiedergabe kenntlich gemacht.

Bielefeld, den 25. März 2004